



University
of Glasgow

Bonnar, Elizabeth-Ann (2005) Attending to visual information for perception and recognition. PhD thesis.

<http://theses.gla.ac.uk/4907/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Attending to Visual Information for
Perception and Recognition

Elizabeth-Ann Bonnar

Department of Psychology
University of Glasgow

Submitted for the degree of PhD to the Higher Degree Committee of the
Faculty of Information and Mathematical Sciences, University of Glasgow

© Elizabeth-Ann Bonnar, 2005

Abstract

A fundamental function of the visual system, and other sensory systems, due to resource limitations, is to optimally select task-relevant information from the barrage of information impinging the retina. In a series of experiments different aspects of attention to visual information for perception and recognition are addressed.

Firstly, what information is selected? In Chapter 2, I begin with addressing the generic problem of knowing which information drives the perception of a stimulus. As a case study example, a first experiment using *Bubbles* (Gosselin and Schyns, 2001) determines the specific information underlying the perception of the stable percepts of an ambiguous image and shows that this information is grounded in different spatial filters processing each image interpretation. A further experiment employs frequency-specific adaptation to induce a perception of the image that is orthogonal to the adapting frequencies, validating this information drives the selective perception of the ambiguous image.

Secondly, if we know the subset of information that is selected for the perception of an input, can the processing of information underlying a percept be selectively suppressed, thereby inducing an alternative percept? In the experiments of Chapter 3, I further apply this spatial frequency adaptation method to test the relevance of *local* image features for the perception of the


ambiguous image, and the recognition of gender in hybrid faces. While the results of the experiment on the perception of the ambiguous image suggest an effective method for testing the role of local stimulus information for perception, the results of the experiment on the recognition of gender in faces showed no effect of adaptation region on perception of gender.

Thirdly, how does selective information use evolve with learning? And what are the mechanism(s) that enable this learning? In Chapter 4, I investigated the evolution of information use in the discrimination of unfamiliar faces using a perceptual learning paradigm. I used the method of noise masking to examine the facial regions observers used, over time, for successful discrimination, and to determine the mechanisms underlying performance improvements. The results showed that the efficiency of observers to use the information available increased with learning differentially across different regions of the face.

Supplementary to examining attention to information in terms of the information content of the stimulus, another series of experiments investigated the effect of attention on the temporal dynamics of processing visual information, and the locus of this effect within the stream of information processing. Using spatial pre-cueing to manipulate attention and a speed-accuracy trade-off (SAT) method to examine the full time-course of visual processing, I analysed, in addition to the behavioural response and the parameters of the SAT function, the lateralised readiness potential (LRP) and other components of the event-related brain potentials (ERPs) to determine the locus of any attentional modulations on the speed of processing. The results of three experiments showed that attention can speed up visual information processing, and that the locus of this effect is at later processing stages related to the categorisation of a stimulus.

Declaration

This thesis has been composed by the undersigned. It has not been accepted in any previous application for a degree. The work, of which this thesis is a record, has been completed by myself, unless otherwise indicated in the text. I further state that no part of this thesis has already been, or is concurrently, submitted for any such degree or qualification at any other university.



Elizabeth-Ann Bonnar

for my mum, June Bonnar

Acknowledgements

The integrity of my supervisor, Hartmut Leuthold, enabled me to complete this thesis: I thank him for his incredible kindness and tolerance, and commitment to scholarship, and for being one of life's good guys.

Philippe Schyns and Frédéric Gosselin were both supervisors/mentors with whom I shared many interesting conversations and warm times. I learned a lot from both of them, and haven't forgotten the good times. Special thanks and acknowledgements are due to Pascal Mamassian and Ines Jentzsch: to Pascal for his integrity, continuous support, encouragement, and eternal friendship, and to Ines for being a patient teacher and friend, and for playing Beethoven's *Moonlight Sonata* so beautifully. Thanks you two.

Lots of good times were spent drinking coffee with Linn Brynildsen, and in pubs with Liza Paul, Becky Champion, Ross Goutcher, Helena Paterson, Wendy Adams, Erich Graf, and latterly, Stephen Butler. Thanks also to the Support staff at Glasgow – particularly Anne Tonner and Lynda Young, and to the Economic and Social Research Council (ESRC) for funding me for three years.

Thanks to Alan McMunnigall for providing a diversion in talking about Joyce, Kelman, Leonard, and inspiring others. And lastly, thanks for support to my mum, June Bonnar, and my aunt, Marie Glazer. It came in many forms. Yes.

Table of Contents

- Abstract 2**
- Acknowledgements..... 5**
- Table of Contents 7**
- List of Figures 11**
- List of Tables..... 15**

- Chapter 1 17**
- Attending to Visual Information for Perception and Recognition: An Introduction..... 17**
 - 1.1 Diagnostic recognition.....18
 - 1.1.1 Task Demands.....20
 - 1.1.2 Available object information.....23
 - 1.1.3 The role of expertise.....26
 - 1.1.4 Summary.....26
 - 1.2 Tools for studying visual information use.....27
 - 1.2.1 Signal Detection Theory.....27
 - 1.2.2 Noise masking.....33
 - 1.2.3 Ideal observer analysis.....35
 - 1.2.4 Reverse correlation.....37
 - 1.2.5 The Bubbles technique.....40
 - 1.2.6 Summary.....42
 - 1.3 Spatial Frequency Channels in Early Vision.....43
 - 1.3.1 Spatial frequency channels in the early visual system.....43
 - 1.3.2 Beyond Gratings: Spatial Frequency Channels and Recognition.....46
 - 1.4 Perceptual learning.....53
 - 1.4.1 The Nature of Perceptual Learning.....54
 - 1.4.2 Potential mechanisms of perceptual learning.....56
 - 1.4.2.1 An early noise model.....59
 - 1.4.2.2 A late noise model.....65
 - 1.4.3 Summary.....71
 - 1.5 Selective spatial attention.....72
 - 1.5.1 Manipulating spatial attention: peripheral and central cueing.....73

- 1.5.2 Early versus late attention effects76
- 1.5.3 ERP studies of selective spatial attention78
 - 1.5.3.1 *Central versus peripheral cueing in ERP studies* 79
- 1.5.5 Measuring the speed of information processing.....84
- 1.5.6 SAT and the Lateralised Readiness Potential (LRP).....89
- 1.5.7 Summary91
- 1.6 Organisation91

Chapter 2 94

The Visual Information Underlying the Perception of a Bi-stable Image..... 94

- 2.1 Revealing the visual information for perception97
- 2.2 Experiment 2.198
 - 2.2.1 Method99
 - 2.2.1.1 *Participants* 99
 - 2.2.1.2 *Stimuli & Apparatus* 99
 - 2.2.1.3 *Procedure*..... 102
 - 2.2.2 Results104
 - 2.2.3 Discussion106
- 2.3 Validating Experiment 2.1 with spatial frequency adaptation.....107
- 2.4 Experiment 2.2109
 - 2.4.1 Method110
 - 2.4.1.1 *Participants* 110
 - 2.4.1.2 *Stimuli & Apparatus* 110
 - 2.4.1.3 *Procedure*..... 112
 - 2.4.2 Results and Discussion.....112
- 2.5 Concluding remarks.....114

Chapter 3116

Perceiving and Recognising Local Image Features116

- 3.1 Experiment 3.1117
 - 3.1.1 Method119
 - 3.1.1.1 *Participants* 119
 - 3.1.1.2 *Stimuli & Apparatus* 119
 - 3.1.1.3 *Procedure and Design*..... 121
 - 3.1.2 Results122
 - 3.1.2.1 *Congruent-adaptation*..... 122
 - 3.1.2.2 *Incongruent-adaptation*..... 124
 - 3.1.3 Discussion and Conclusions124
- 3.2 Locally adapting information for recognition126
- 3.3 Experiment 3.2127
 - 3.3.1 Method128
 - 3.3.1.1 *Participants* 128
 - 3.3.1.2 *Stimuli & Apparatus* 129
 - 3.3.1.2.1 *Diagnostic Hybrids*..... 129

3.3.1.2.2 Adaptation Stimuli.....	130
3.3.1.3 Procedure and Design.....	132
3.3.2 Results	133
3.3.2.1 Diagnostic-adaptation condition	133
3.3.2.2 Non-diagnostic adaptation condition.....	134
3.3.3 Discussion	135
3.3.3 Concluding remarks	137

Chapter 4138

Learning to Discriminate Features in Noise138

4.1 Mechanisms of perceptual learning.....	139
4.2 Experiment 4.1	141
4.2.1 Methods.....	142
4.2.1.1 Participants	142
4.2.1.2 Stimuli	142
4.2.1.3 Procedure.....	144
4.2.1.3.1 Human Observers	144
4.2.1.3.2 Ideal Observer.....	145
4.2.1.4 Data Analysis	148
4.2.2 Results	149
4.2.3 Discussion	159
4.2.4 Concluding remarks	163

Chapter 5164

Attention Modulates the Speed of Visual Information Processing164

5.1 Determining the locus of a temporal modulation of attention	165
5.2 Experiment 5.1	168
5.2.1 Method	169
5.2.1.1 Participants	169
5.2.1.2 Stimuli & apparatus	169
5.2.1.3 Procedure.....	171
5.2.1.4 Electrophysiological Recordings.....	172
5.2.1.5 Data Analysis	174
5.2.1.5.1 Behavioural data analysis	174
5.2.1.5.2 EEG data analysis	176
5.2.2 Results	177
5.2.2.1 Behavioural measures	177
5.2.2.1.1 Reaction time and Accuracy	177
5.2.2.1.2 Speed-Accuracy Tradeoff.....	179
5.2.2.2 Electrophysiological measures.....	181
5.2.2.2.1 The Stimulus-locked LRP (S-LRP).....	181
5.2.2.2.2 S-LRP onset latency as a chronometric measure to fit SAT function	184
5.2.2.3 The Response-locked LRP (LRP-R).....	186

5.2.2.4 Event-related potentials (ERPs).....	189
5.2.3 Discussion	198
5.3 Experiment 5.2	201
5.3.1 Method	202
5.3.1.1 Participants	202
5.3.1.2 Stimuli and Procedure.....	202
5.3.2 Results	203
5.3.2.1 Behavioural data	203
5.3.2.1.1 Reaction time and accuracy	203
5.3.2.1.2 Speed-Accuracy Trade-off.....	204
5.3.2.2 Electrophysiological Measures.....	206
5.3.2.2.1 The S-LRP	206
5.3.2.2.2 The LRP-R.....	210
5.3.2.2.3 Event-related potentials	211
5.3.3 Discussion	220
5.4 Experiment 5.3	221
5.4.1 Method	222
5.4.1.1 Participants	222
5.4.1.2 Stimuli and Procedure.....	222
5.4.2 Results	223
5.4.2.1 Behavioural data	223
5.4.2.1.1. Reaction time and accuracy	223
5.4.2.1.2. Speed-Accuracy Trade-off.....	224
5.4.2.2 Electrophysiological measures.....	226
5.4.2.2.1 The S-LRP.....	226
5.4.2.2.2 S-LRP onset latency as a chronometric measure to fit SAT function	229
5.4.2.2 The LRP-R.....	231
5.4.2.3 The effect of cueing on ERPs.....	234
5.4.3 Discussion	241
5.5 Concluding remarks.....	242
Chapter 6	244
General Discussion	244
6.1 Summary	244
6.2 Issues, methodological and theoretical.....	250
6.2.1 Does <i>Bubbles</i> induce atypical observer strategies?.....	250
6.2.2 The concept of diagnosticity.....	252
6.2.3 Determining the information selected for perception	253
6.3 Conclusion.....	254
References	255

List of Figures

Figure 1.1. The Diagnostic Recognition Framework (adapted from Schyns, 1998).
.....19

Figure 1.2. Signal Detection Theory Response Distributions for a simple
detection task.....31

Figure 1.3. The technique of noise masking.....34

Figure 1.4. Examples of stimuli created from methodological techniques available
to manipulate visual information.....40

Figure 1.5. The Contrast Sensitivity Function (CSF) of an adult human observer.
.....44

Figure 1.6. An example of spatial scales47

Figure 1.7. Hypothetical noise masking functions for three potential mechanisms
of perceptual learning.....58

Figure 1.8. Pelli’s black-box model of a human observer in a perceptual task....60

Figure 1.9. Noise masking functions from Gold et al. (1999b, 2004).....63

Figure 1.10. The Perceptual Template Model (PTM) of Doshier and Lu.....65

Figure 1.11. Noise masking functions from Lu and Doshier (1998).....68

Figure 1.12. A hypothetical speed-accuracy trade-off function86

Figure 1.13. Hypothetical SAT functions..87

Figure 1.14. The results of Carrasco and McElree (2001)89

Figure 2.1. Salvador Dali’s *Slave Market with the Disappearing Bust of Voltaire*
(1940).....95

Figure 2.2. The application of Bubbles to the 3D-image generation space101

Figure 2.3. Computing the significance of each region in driving each percept.	103
Figure 2.4. The information driving the stable percepts of Dali's <i>Slave Market with the Disappearing Bust of Voltaire</i>	105
Figure 2.5. Stimuli and design of Experiment 2.2.....	111
Figure 3.1. The stimuli and design of Experiment 3.1.	121
Figure 3.2. The results of Experiment 3.1.....	123
Figure 3.3. Stimulus generation for Experiment 3.2.	131
Figure 3.4. Stimuli and design of Experiment 3.2.....	133
Figure 3.5. The results of Experiment 3.2.....	134
Figure 4.1. The stimulus generation process.	144
Figure 4.2. Example of a single trial in Experiment 4.1.	145
Figure 4.3. The confusion matrix for the image set used in Experiment 4.1.....	147
Figure 4.4. Confusion matrices for each of the different face regions, derived from the ten faces used in Experiment 4.1..	148
Figure 4.5. Cumulative Gaussian functions plotting proportion correct as a function of signal energy contrast for Observer D.K. in the face discrimination task of Experiment 4.1.	151
Figure 4.6. The results of Observer V.L. depicted as contrast threshold versus noise image maps.....	155
Figure 4.7. Calculation efficiency measures for Observer S.H.....	157
Figure 4.8. Calculation efficiency measures for Observer D.K.....	158
Figure 4.9. Calculation efficiency measures for Observer V.L.	159
Figure 5.1. The stream of information processing and the electrophysiological markers of processing stages.....	167
Figure 5.2. A typical trial sequence of Experiment 5.1.....	171
Figure 5.3. Electrode layout for the Biosemi 64 channel system.....	173
Figure 5.4. Average time course data for Experiment 5.1.....	179
Figure 5.5. The stimulus-locked LRP and HEOG for Experiment 5.1.....	183

Figure 5.6. Average discrimination accuracy (d' units) as a function of S-LRP latency for Experiment 5.1.	186
Figure 5.7. The response-locked LRP and HEOG for Experiment 5.1	188
Figure 5.8. Grand-averaged waveforms for cued and neutral conditions of Experiment 5.1.....	191
Figure 5.9. Mean amplitude (measured in μ V) at electrodes PO7, PO8, PO3, PO4, O1, O2, and POZ in the time interval 170-200 ms post-cue for the cued and neutral conditions of Experiment 5.1.	194
Figure 5.10. Difference waveforms for Cued - Neutral conditions of Experiment 5.1 at each response lag for seven electrode sites (PO7, PO8, POZ, PO3, PO4, O1, and O2).	195
Figure 5.11. Scalp topographic voltage maps of the difference between cued and neutral trials from 90 ms to 370 ms.	196
Figure 5.12. Grand-averaged waveform depicting the P3 component for cued and neutral trials at electrode site Pz.	197
Figure 5.13. Scalp topographic voltage maps of P3 activation for cued (top) and neutral (bottom) trials	198
Figure 5.14. Average time course data for Experiment 5.2.....	205
Figure 5.15. The stimulus-locked LRP and HEOG for Experiment 5.2	207
Figure 5.16. The response-locked LRP and HEOG for Experiment 5.2	209
Figure 5.17. Grand-averaged waveforms for Experiment 5.2 for valid and invalid conditions..	213
Figure 5.18. Mean amplitude for valid and invalid conditions at lateral parietal-occipital PO7 and PO8, PO3 and PO4, lateral occipital O1 and O2, and the midline parietal-occipital POZ, at time interval 260-320 ms.	215
Figure 5.19. Difference waveforms between valid and invalid trials when the target stimulus was presented along the vertical meridian (collapsed across top and bottom locations) and along the horizontal meridian (collapsed across left and right locations).	217
Figure 5.20. Scalp topographic voltage maps for valid minus invalid cues averaged over top and bottom target locations in the top picture and over left and right target locations in the bottom picture.	218
Figure 5.21. Scalp topographic voltage map of P3 activation in Experiment 5.2..	219

Figure 5.22. Average time course data for Experiment 5.3.....	225
Figure 5.23. The stimulus-locked LRP and HEOG for Experiment 5.3.	228
Figure 5.24. Average S-LRP onset data for Experiment 5.3.....	230
Figure 5.25. The response-locked LRP and HEOG for Experiment 5.3.....	233
Figure 5.26. Grand-averaged waveforms for Experiment 5.3 for valid and invalid conditions at electrodes PO7 PO8, POZ, PO3, PO4, O1, and O2.	235
Figure 5.27. Mean amplitude for valid and invalid conditions at lateral parietal-occipital PO7 and PO8, PO3 and PO4, lateral occipital O1 and O2, and the midline parietal-occipital POZ, at time interval 460-520 ms (around 340-400 ms following target onset).	237
Figure 5.28. Difference waves (valid(ERP) minus invalid(ERP)) for the six response lags of Experiment 5.3.....	238
Figure 5.29. Topographical maps of the difference activation between valid and invalid (valid minus invalid) trials of Experiment 5.3.	239
Figure 5.30. Grand-averaged ERP waveform showing peak P3 latency for valid and invalid trials at electrode Pz.....	240
Figure 5.31. Scalp topographic voltage maps of P3 peak latency for valid (left) and invalid (right) trials, at 440 ms and 470 ms, respectively, for Experiment 5.3.	240

List of Tables

Table 4.1. Contrast thresholds for Observer D.K	152
Table 4.2. Contrast thresholds for Observer S.H.....	153
Table 4.3. Contrast thresholds for Observer V.L	154
Table 5.1. Mean reaction times (ms) and accuracy (proportion correct) for cued and neutral conditions at six response lags (SD in parentheses).	178
Table 5.2. The mean onset-latency (ms) of the stimulus-locked LRP (S-LRP) for cued and neutral trials at six response lags in Experiment 5.1.....	182
Table 5.3. The mean onset-latency (ms) of the response-locked LRP (LRP-R) for cued and neutral trials at six response lags in Experiment 5.1 (standard deviations in parentheses).....	187
Table 5.4. Greenhouse-Geisser corrected <i>F</i> -values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data for Experiment 5.1 over electrode sites PO7, PO8, PO3, PO4, O1, O2, and POZ ($\dagger p > 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.005$).....	190
Table 5.5. Greenhouse-Geisser corrected <i>F</i> -values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on the ERP amplitude data for Experiment 5.1 on seven electrode sites: occipital electrodes O1 and O2; parietal-occipital electrodes PO3, PO4, PO7 PO8, and the midline electrode POZ. ($\dagger p > 0.05$, $*p < 0.05$, $**p < 0.01$).....	193
Table 5.6. Mean reaction times and performance accuracy for valid and invalid conditions at six response lags.	204
Table 5.7. The mean onset-latency (ms) of the stimulus-locked LRP (S-LRP) for the valid and invalid conditions at six response lags in Experiment 5.2 (standard deviations shown in parentheses).....	206

Table 5.8. The mean onset-latency (ms) of the response-locked LRP for valid and invalid trials at six response lags of Experiment 5.2 (SD in parentheses).....210

Table 5.9. Greenhouse-Geisser corrected *F*-values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data for Experiment 5.2 ($\dagger p>0.10$, $*p<0.05$, $***p<0.001$).....212

Table 5.10. Greenhouse-Geisser corrected *F*-values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude of seven electrodes ($\dagger p>0.10$, $**p<0.01$, $***p<0.001$).214

Table 5.11. Mean reaction times and performance accuracy for valid and invalid conditions at six response lags (SD in parentheses).....224

Table 5.12. The mean onset-latency (ms) of the stimulus-locked LRP for valid and invalid trials at six response lags of Experiment 5.3.227

Table 5.13. The mean onset-latencies (ms) of the response-locked LRP for valid and invalid trials at six response lags of Experiment 5.3.232

Table 5.14. Greenhouse-Geisser corrected *F*-values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data for Experiment 5.3 ($\dagger p>0.10$, $*p<0.05$, $***p<0.001$).....234

Table 5.15. Greenhouse-Geisser corrected *F*-values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data of Experiment 5.3 for seven electrode sites ($\dagger p>0.10$, $*p<0.05$)..236

Chapter 1

Attending to Visual Information for Perception and Recognition: An Introduction

Phenomenologically, our perceptual experience of the world is coherent and rich in detail. Yet, there is overwhelming converging evidence suggesting that our perception of the world is, by necessity, highly selective in nature, indicating in reality we attend only to a subset of the sensory information available. This ability to select, and suppress, information is immensely valuable to an observer in a dynamic environment. Empirical support for the selective nature of attention derives from early studies examining eye movements in face and scene perception that depicted eye scan paths involving a series of fixations over only a portion of the information available (Yarbus, 1967); a finding that has been extended over decades. Additionally, studies of change blindness, that is the failure to detect changes in a scene, offer compelling, and sometimes surprising demonstrations of the extent to which external events are neglected (for reviews see Rensink, 2002, and Simons and Rensink, 2005).

This thesis is concerned with a range of issues related to the selective use of visual information for perception and recognition. To set the scene for the empirical investigations described in Chapters 2 to 5, in this chapter, I review previous research related to questions of selective attention to visual information. Firstly, in Section 1.1, I will consider a theoretical framework, the Diagnostic Recognition Framework (Schyns, 1998) that has been proposed in which to frame questions of information use for perception and recognition. Next, in Section 1.2, I introduce some of the methodological tools that are available to study questions of information use and the different research areas in which they have been employed. Subsequent sections review research focussing on different aspects of visual information use. Section 1.3 reviews the literature related to the role of spatial frequency channels in perception and recognition, and in Section 1.4, I consider how visual information use may evolve with perceptual learning, and describe mechanisms that have been proposed to enable learning. Lastly, in Section 1.5, I describe behavioural and electrophysiological research on selective spatial attention to assess the functions of attention and the processing stages modulated by attention.

1.1 Diagnostic recognition

On acknowledging that observers, at any given time, attend only to a subset of information available in the environment, the question arises pertaining to the determinants of this selection; for a perceptual or recognition task which information should the observer select? For many years the question of information use suffered neglect; researchers studying recognition of faces,

objects and scenes, commonly focussed on the structure of memory representations in explaining recognition phenomena (Liu, Kersten, and Knill, 1999; Schyns, 1998). While debates regarding the nature of object representations, in explaining recognition, sustained (e.g. viewpoint-dependence versus independence, see Tarr and Bülthoff, 1995 and Biederman and Gerhardstein, 1995, respectively for insight into this debate), analyses of the information available in the input, and consideration of task demands suffered neglect.

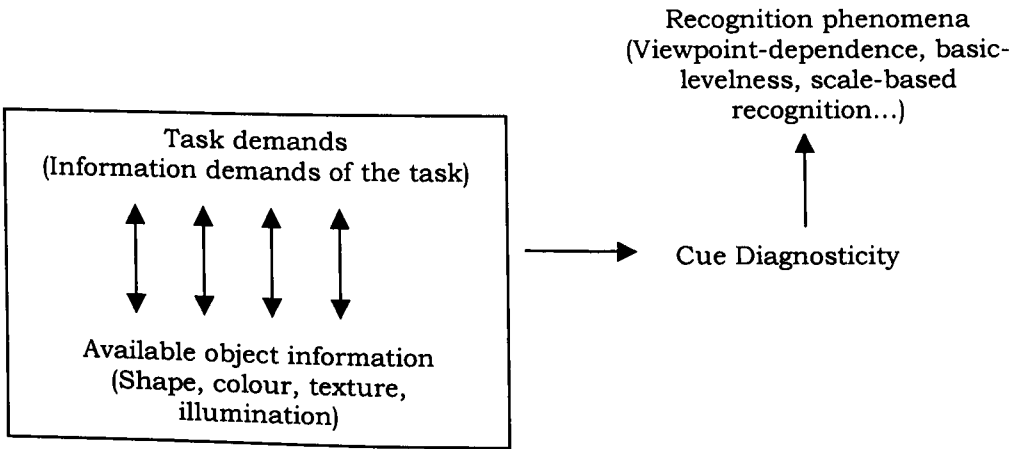


Figure 1.1. The Diagnostic Recognition Framework (adapted from Schyns, 1998). The framework highlights how the interaction between the information demands of a task and the information available jointly determine the subset of information that becomes diagnostic to resolve a task. According to this framework, information diagnosticity could explain the image information subtending various recognition phenomena.

The Diagnostic Recognition Framework (see Figure 1.1), articulated by Schyns (1998), attempts to redress the balance between the information

represented in memory, the information available to the observer as input, and the constraints imposed by the visual and cognitive systems in extracting this information. Schyns (1998) proposes that the interactions between the information demands of a recognition task and the availability of useful object information jointly determine the subset of information selected by the observer, in other words, the information *diagnostic* to resolve a visual task. Still, for a given task, how do we determine which information is the diagnostic information? Further still, how can we measure the availability of *useful* object information? The advent of sophisticated methodological tools and their application to the study of information use in visual recognition is enabling these questions to be addressed, enhancing knowledge of the information observers use to resolve daily visual tasks.

To understand the Diagnostic Recognition Framework as a context in which to consider aspects of information use for perception and recognition, I first consider how task demands, and their interaction with the availability of object information, might modulate information use. A description of the research methods employed to study questions of information use for perception and recognition will follow in Section 1.2.

1.1.1 Task Demands

Imagine a single human face as present visual input, multiple classifications are possible – this face could be male, young, sad, familiar, and his identity, your brother Mark. Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976) indicated that a single object might be categorised at three different levels of specificity -

the superordinate (e.g. tree, musical instrument), basic (e.g. birch, guitar), and subordinate (e.g. White birch, Folk guitar) levels. As a determinant of information selection, task demands refer to the information constraints imposed by these multiple categorisations that are possible in recognising a single visual input. The information demands of the recognition task, for example in judging the gender or expression or identity of this single face, differentially modulate the information attended in that visual input. Task demands, therefore, necessarily reflect the information represented in memory for the particular categorisation or recognition of a given visual input.

Task demands have been shown to modulate the selection of visual input at early processing stages within the visual system (Oliva and Schyns, 1997). A fundamental aspect of the perception and recognition of any visual input is the spatial scales at which an input is represented. The spatial scale of an input, as viewed by an observer, is determined by the size of the input and its distance from the observer. Consider again as an example a human face as input; from a distance you are unable to perceive the fine details of the face, such as the eyes and shape of the mouth, what you see instead is a coarse outline of the face with these finer details becoming available only as the face comes nearer. Different features of the face, and of other visual inputs therefore, are represented at different spatial scales, from fine to coarse. A classical finding in vision research is the existence of a set of channels in the early visual system that are responsible for processing visual input at these different spatial scales (Blakemore and Campbell, 1969a, 1969b; see Section 1.3). Studies investigating the effect of categorisation task demands on the selective use of visual information have demonstrated a differential use of this spatial scale information

in studies of scene (Oliva and Schyns, 1997) and face (Schyns, Bonnar, and Gosselin, 2002; Schyns and Oliva, 1999) categorisations. Moreover, Schyns et al. (2002) showed that this information also changed as a function of the particular face categorisation task. Specifically, the 2D facial information observers attended to resolve one of three face categorisation tasks, identity, gender or expression, showed that the mouth region was selected at each spatial scale for the identity and expression tasks, whereas for the gender task this region was entirely neglected at all scales. For the identity task, both eyes were selected at all spatial scales, whereas for the gender task, one eye was sufficient (from the observer's viewpoint, the left eye), and in the expression task both eyes were neglected.

To resolve tasks of scene and face categorisation, these studies showed that the information observers attended was determined by the presence of task-dependent, diagnostic information that differed across spatial scales both within and across different tasks. Since only the categorisation task itself differed between observers (all participants experienced the same training procedures and were presented with identical face stimuli), it was concluded that the presence of diagnostic cues determined the information selected for each categorisation. Task demands, therefore, form an important contributing factor in the selection of visual information for perception and recognition. Importantly, the findings of these studies are not restricted to the study of face and scene recognition but demonstrate an important finding regarding the architecture of the visual and cognitive systems, specifically that high level task demands modulate the low level processing of visual input (see Pylyshyn, 1999 for a thesis against this idea).

Next, I consider how the availability of information determines the information diagnostic to resolve a task.

1.1.2 Available object information

The availability of object information for perception and recognition is constrained by factors internal, and external, to the observer. Internal constraints relate to processing inefficiencies; the neural wiring of the visual system imposes limitations on the spatial and temporal resolution in which object cues may be detected or discriminated. Limitations on the availability of visual information external to the observer non-exclusively include degraded lighting conditions, occlusion, spatial scale, crowding, translation and rotation. Together, these factors limit the information theoretically available to the observer. Additionally, objects that share category membership often comprise highly similar features; faces constitute two eyes, a nose, and a mouth that are arranged fairly homogenously: how then are we able to effortlessly discriminate thousands of faces? What is the distinguishing information that is available in the input to enable these fine-detailed discriminations? To achieve a full account of the selective use of information for perception and recognition, an examination of the information available in the input is crucial; a dismissal of this factor has led, sometimes, to a misinterpretation of recognition phenomena. For example, the phenomena of viewpoint-based recognition, the observation that object recognition performance is often enhanced when an object is viewed from specific viewpoints, is one such area of research that in neglecting the factors of task constraints and information availability led to a misinterpretation of the phenomena under observation (Schyns, 1998). A brief summary of the debate regarding viewpoint-based recognition follows.

In the study of object recognition, a debate arose in the late 1980's to mid 1990's regarding the viewpoint-invariance of object recognition. Although object recognition is relatively invariant to changes in viewpoint, observers do exhibit increases in accuracy and reaction time for categorising objects at specific viewpoints, and decreases in these performance measures as an object is increasingly oriented away from this viewpoint (Tarr and Bülthoff, 1995). The interpretation of this phenomenon focussed on determining the structure of object representations in memory. The 'viewpoint-dependent' proponents interpreted this phenomenon as evidence that objects are represented in memory as a series of multiple viewpoints (Tarr and Bülthoff, 1995); if performance improved for a particular viewpoint this was interpreted as evidence that this viewpoint was stored in memory. On the other side of the theoretical divide, the 'viewpoint-invariant' theorists Biederman and Gerhardstein (1993, 1995) suggest that object recognition is viewpoint-independent. Instead, they argue that objects are represented in memory as configurations of volumetric primitives, called geons. The evidence suggesting viewpoint dependency, by these theorists, is attributed to visual processes external to those responsible for normal object recognition.

In this debate, while focussing on the structure of object representations in memory, researchers neglected to consider how task demands and the availability of object information changed with variations in viewpoint. The details of this debate are too lengthy to be discussed here and are of no special relevance other than to illustrate the importance of considering how task demands and the information in the input interact to effect processing (cf. Schyns, 1998 for a summary of this debate). For example, under the diagnostic recognition

framework, the cues that are diagnostic for object recognition will vary across different views for different categorisation tasks; these cues may or may not be available within a given viewpoint. The unavailability of some cues in a given viewpoint may impair categorisation performance for one type of task, and not for another as this cue may not be diagnostic for this task. The point to be emphasised is that to attribute behavioural performance to a visual input, and to subsequently infer the representational format of this input in memory, one must firstly have knowledge about the informational content of the input and its availability to the observer in matching with its memory representation (Liu, 1996).

The emergences of studies in object recognition that explicitly measure the information content of a stimulus, and its availability, have provided a fuller account of recognition. In relation to the debate regarding viewpoint-invariance versus viewpoint-dependency of object recognition, these studies have demonstrated that an object's viewpoint-invariance or dependency is determined by the information complexity contained within a specific view and the extent to which it matches with the object's representation in memory (Liu, 1996; Liu, Knill, and Kersten, 1995, 1999). Tjan and Legge (1998) developed a measure to quantify the viewpoint complexity of an object recognition task using an ideal observer approach (the concept of an ideal observer is described in Section 1.2.3). This work emphasised the importance of measuring the information content of any visual input, the demands of the task, and stimulus set used. As a result, this work significantly evolved the debate in object recognition, specifically, regarding the interpretation of object recognition phenomena such as viewpoint-

dependency, and generally in highlighting the importance of understanding the informational content of a visual input in any perceptual task.

1.1.3 The role of expertise

An additional factor, not explicitly considered in the Diagnostic Recognition Framework is the effect of previous experience with a visual input, and the effects of this experience in selecting a subset of information to detect or discriminate this input. The level of familiarity, held by an observer, for a particular visual input, may modulate the selection of information used to resolve a visual task. Indeed studies investigating the effects of perceptual learning show that information use does change over time (Gold, Sekuler, and Bennett, 2004).

1.1.4 Summary

An approach to information use that focuses only on the information represented in memory, or only to the information in the input will fail in capturing the incredible capacity to perceive and recognise the seemingly infinite number of objects impinging the retina. The interaction between the information represented in memory for a particular categorisation, and the information available in the input determine the subset of information that earns a diagnostic status, the subset of information to resolve the task at hand. Also relevant, as a determinant of selective information use, is the level of familiarity held by an observer for a particular class of objects.

The availability of research methodologies to approach questions of information use has initiated research with diverse populations, including clinical

patient groups with brain damage inducing impairments in judging emotional expressions, identifying faces, hemi-spatial neglect, as well as looking at how autistic children process faces. These methodologies will now be described.

1.2 Tools for studying visual information use

In recent years methodological developments have enabled finer analyses of the information available to observers for a visual task, and in determining the subset of information that is used to resolve the task. Methods such as *Bubbles* (Gosselin and Schyns, 2001), the construction of hybrid stimuli (Oliva and Schyns, 1997; Schyns and Oliva, 1999), reverse correlation (Gold, Murray, Bennett, and Sekuler, 2000), noise masking (Gold, Bennett, and Sekuler, 1999b; Lu and Doshier, 1998; Solomon and Pelli, 1994), and ideal observer analyses (Liu et al., 1995; Tjan and Legge, 1998) have been employed to probe problems of information use.

Prior to describing these methods as tools to study questions of information use, for clarity's sake, an overview of the general framework on which some of these techniques are grounded, namely signal detection theory (SDT), will preface summaries of these techniques. Extensive focus will be given to methods employed in the chapters to follow.

1.2.1 Signal Detection Theory

The methods of noise masking, ideal observer analyses, and reverse correlation appeal explicitly to the concepts of signal detection theory (Green and Swets, 1966), while the bubbles method borrows some of its concepts. Therefore, for the purposes of clarity, an introduction to SDT follows.

Signal detection theory (SDT) provides a general framework in which to measure sensory/perceptual processes, taking into account both the sensitivity of an observer to a sensory event and the subsequent decision processes relating to the event. A central tenet of SDT is the proposal that detection decisions result from an initial sensory stage whose outputs feed into a decision stage, from which the outcome is a decision regarding the occurrence of a sensory event. A key assumption of SDT is that the strength of all sensory events is continuous, so that over a series of trials there is a distribution of responses for a signal event, and a distribution for noise only. SDT states that all decision processes occur in the presence of uncertainty so that an observer's internal responses to any sensory event are, as a result, probabilistic.

To illustrate the uncertainty associated with a detection task facing the cognitive system, consider the following scenario. Imagine the radiologist's task of examining a medical X-ray image for potential abnormalities. The complexity of the image makes it difficult to discriminate normal and abnormal features, and the ability to do so accurately is dependent on both the observer's visual *sensitivity* and *decision-criterion*. The measurement of an observer's sensitivity characterises the processing *efficiency* of the perceptual system to a sensory event, whereas an observer's decision-criterion relates to the observer's subjective response to a sensory event. Due to the difficulty of this task and subsequent uncertainty, several outcomes are possible. For example, an abnormality may be present and our radiologist may *accurately* detect its presence (in SDT, this response is referred to as a 'hit'), or the abnormality may be absent and its absence may be *accurately* detected (termed a 'correct rejection'). These responses are the desired outcomes, a signal (abnormality) is detected when it is

present, and a signal is declared absent when it is absent. The radiologist is doing well! However, two other outcomes are also possible: the radiologist may *falsely* detect the presence of an abnormality when there is in fact none (referred to as a 'false alarm'), or may *falsely* detect an absence of an abnormality when it is actually present (termed a 'miss'). Clearly, in the face of uncertainty, any observer, and perhaps particularly our radiologist, wishes to minimise the number of 'false alarms' and 'misses', and sets their decision-criterion accordingly.

The observer's visual sensitivity depends on the strength of any abnormality in the image, and on the observer's experience in examining such images (Sowden Davies, and Roling, 2000). However, the human observer is also limited by other factors, for example, the optics of the eye results in stimulus information loss prior to transmission through the visual system, and internal noise within the visual system, generated by random neural firing naturally associated with processing, also constrains performance. Additionally, the mechanisms with which human observers encode visual input may not correspond perfectly to the spatial and/or temporal distribution of the input, and once encoded the observer may employ less than ideal decision processes. All of these factors influence the observer's efficiency in using the sensory information available.

In relation to decision-criterion, multiple factors influence an observer's decision that an event has occurred, including the expectations of the observer or pay-offs, each contributing to the probability of detection. In setting their decision-criterion, that is, in forming a decision regarding the presence or absence of any signal, an observer will take into account the consequences of making an inaccurate decision. Turning back to the radiologist example, one

radiologist may be extremely conservative in declaring the presence of an abnormality in the face of uncertainty; they may fear that the cost of an unnecessary biopsy is too high, and therefore set their response-criterion towards signal absent decisions. Conversely, a second radiologist may be less conservative, believing that early detection is paramount, and may therefore set their response-criterion towards signal present decisions. The overlap between the *signal present* response distribution and *signal absent* response distribution is depicted in Figure 1.2. This overlap in response distributions between detecting the presence or absence of a signal causes a problem for the observer: Which response to choose? Fortunately, SDT uniquely provides a formal description of how an observer's visual sensitivity and decision-criterion determine the percentage of hits, misses, false alarms and correct rejections on visual detection and discrimination tasks. Figure 1.2 indicates the areas under these distributions where these response categories fall.

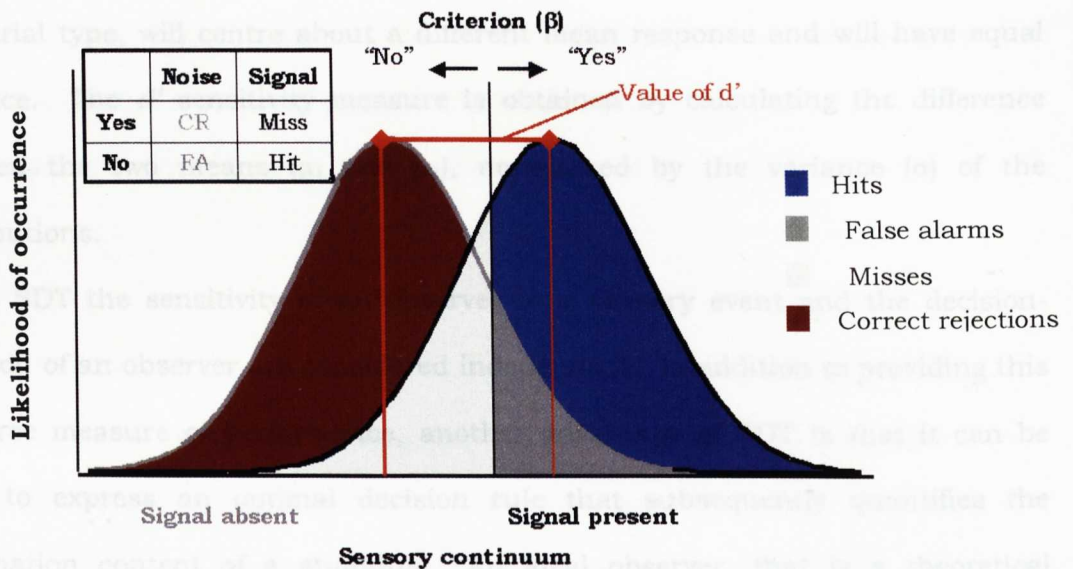


Figure 1.2. Signal Detection Theory Response Distributions for a simple detection task. The horizontal axis represents the observer's level of certainty that a signal is present and the vertical axis, the probability that response will occur. The normal distribution on the left represents the distribution of internal responses when the signal is absent, while the distribution on the right represents the distribution of responses when the signal is present. Notice the overlap between these two distributions; the internal response on signal absent trials may exceed the internal response for a signal present trial incurring detection errors. The distance between the means of these two distributions indicates the sensitivity (measured in d' units) of the observer. The black vertical line represents the decision criterion of the observer (β).

A detection task would involve presentation, on alternate trials, a signal embedded in noise, or noise alone, whereas in a discrimination task the presentation would entail one of at least two unique signals embedded in noise. For a detection task (Figure 1.2 depicts the case for a detection task), there will be a distribution of internal responses for each trial type, a distribution for noise only trials, and one for signal plus noise trials, which is referred to as signal distribution. For the discrimination task, there will be a distribution of responses for each signal in a 1-of- n discrimination task (for a task involving two signals, S_1

and S_2 , there will be two distributions). The internal response distribution, for each trial type, will centre about a different mean response and will have equal variance. The d' sensitivity measure is obtained by calculating the difference between the two means (μ_1 and μ_2), normalised by the variance (σ) of the distributions.

In SDT the sensitivity of an observer to a sensory event and the decision-criterion of an observer are considered independent. In addition to providing this bias-free measure of performance, another advantage of SDT is that it can be used to express an optimal decision rule that subsequently quantifies the information content of a stimulus. An ideal observer, that is a theoretical observer that makes optimal use of all of the available stimulus information, can be used to obtain a benchmark of human performance. A comparison of the performance of this ideal observer with a human observer provides a measure of the processing efficiency of the human observer in using the visual information available.

In comparing an observer's performance on a visual task to either different object cues, or to the same object over time, any change in the variability of the internal noise, and thus the sensitivity of the observer, would be evident in the variance of the distributions. An alteration in sensitivity would also be evidenced in the distance between the means of the internal response distributions, reflecting a change in the efficiency of the observer, that is, how efficiently the observer uses the information available. An observer's internal noise and their efficiency in using the information available, therefore, determine sensitivity.

The provisions of independent measures of sensory sensitivity and response strategies, as well as the comparative tool of employing ideal observer analysis,

have combined in SDT to provide a popular framework for researching visual perception and identification processes. Recently, promising methods grounded in SDT have been developed to yield fruitful insights into visual information use in perception and recognition. It is to these methods we now turn.

1.2.2 Noise masking

The method of noise masking as a tool to study visual perception and recognition is a variant of techniques employed in engineering to measure the inherent noise in a system by introducing a quantity of external noise. To distinguish between the components determining visual sensitivity, that is, internal noise and efficiency, Pelli (1990) suggested that an observer's internal noise, unknown to the experimenter, could be measured by referring to the quantity of experimenter-added external noise in a stimulus. Procedurally, to disentangle the contribution of an observer's internal noise and efficiency to their overall sensitivity to a visual stimulus, detection or discrimination thresholds (i.e. contrast thresholds) are measured in various densities of externally added noise; Figure 1.3 depicts stimuli masked in experimentally-added noise, low and high densities, respectively. These thresholds are then plotted in noise-masking functions that demonstrate the effect on an observer's contrast threshold of adding increasing levels of external noise; thresholds increase as noise density increases (Pelli, 1990). The plot in Figure 1.3 depicts a noise-masking function; note at low levels of external noise there is little effect on contrast thresholds, yet at higher levels of noise density thresholds increase. These increases are evident only when the external noise exceeds the magnitude of an observer's intrinsic

internal noise. This ‘kink point’ of the noise-masking function denotes the *equivalent input noise*, that is, the estimate of the observer’s internal noise. A measure of this equivalent internal noise is obtained by adding a density of external noise to the stimulus that will exceed the magnitude of an observer’s internal noise two-fold (Pelli and Farrell, 1999).

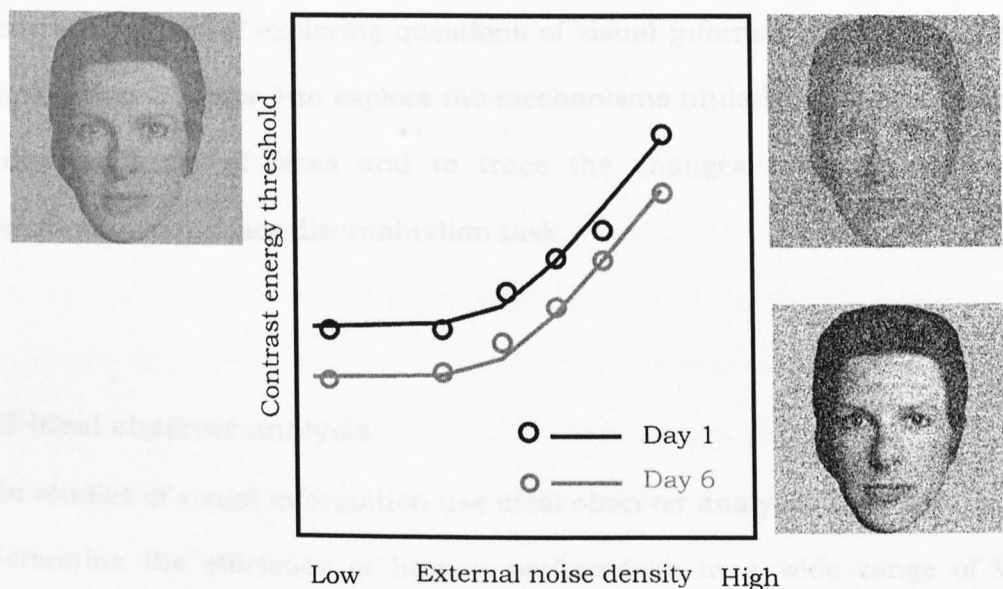


Figure 1.3. The technique of noise masking. The figure depicts a hypothetical noise-masking function with external noise density on the horizontal axis (increasing in density from left-to-right), and contrast energy threshold on the vertical axis. The accompanying pictures depict face stimuli embedded in Gaussian noise. The left image depicts a face of mid-contrast embedded in a low level of Gaussian noise. The right face images depict a face in a higher level of Gaussian noise, of the same contrast in the top picture, and of a higher contrast in the bottom picture. The function illustrates how contrast energy thresholds increase as external noise density increases, and the possible change in contrast thresholds at different noise levels as a function of learning.

This method of noise masking can also be used to compare human performance with that of an ideal observer, thus providing a measure of processing efficiency. The method has been adapted to address diverse questions in the visual domain, including characterising the channel underlying letter recognition (Solomon and Pelli, 1994), determining the mechanisms enabling performance improvements in selective spatial attention tasks (Lu and Doshier, 1998) and the mechanisms of perceptual learning (Doshier and Lu, 1998, 1999; Gold et al., 1999b; Gold, Sekuler, and Bennett, 2004; Lu and Doshier, 2004). In the current context of exploring questions of visual information use, the method is employed in Chapter 4 to explore the mechanisms underlying improvements in the discrimination of faces and to trace the changes, over time, of visual information use in a face discrimination task.

1.2.3 Ideal observer analysis

In studies of visual information use ideal observer analyses have been applied to determine the efficiency of human performance in a wide range of visual perception and recognition tasks (Barlow, 1980), from motion perception (Wallace and Mamassian, 2003) to 3D objects (Liu, Knill, and Kersten, 1995; Tjan, Braje, Legge, and Kersten, 1995), letters (Gold, Bennett, and Sekuler, 1999a; Parish and Sperling, 1991; Solomon and Pelli, 1994) and faces (Gold et al., 1999a). The efficiency measure, that is how effectively the human observer uses the information available, is computed as the ratio of human sensitivity, measured in d' units, to that of the ideal observer. The theoretical ideal observer uses all of the available stimulus information to resolve a task optimally. Comparing human

performance to that of an ideal observer therefore provides an upper limit of the stimulus information available to the human observer to resolve a perceptual task. For example, in a task of face discrimination, in which there are ten face signals each of 256 x 256 pixels, the ideal observer can use the information represented at each pixel that discriminates each signal from the others. An optimal strategy in this type of task is to cross-correlate the input signal, for example a face, with each of the other faces in memory and select that which yields the highest correlation. The ideal observer's performance on this task is limited only by the similarity between the ten face signals, and to prevent ceiling performance, a noise source added by the experimenter. In contrast, the performance of the human observer is limited by an internal noise source, and less-than-optimal strategies. For example, the human observer may not focus on those regions that show the highest variance between the input image and the other images in memory, but may instead focus on features such as the eyes and mouth that to the human observer are diagnostic to resolve the task.

Gold et al. (1999b, 2004) utilised an ideal observer in a series of experiments aiming to characterise the mechanisms underlying perceptual learning. Using the techniques of noise masking and ideal observer analysis to determine the mechanisms enabling performance improvements over time, Gold et al. measured the equivalent input noise and efficiency of observers in face and texture discrimination tasks. Combining the techniques of noise masking and ideal observer analysis enabled a distinction between a perceptual learning mechanism functioning to reduce internal noise, as opposed to a mechanism that operates to increase processing efficiency. The ideal observer utilised by Gold et al. was an observer that on presentation of a stimulus (a signal masked by noise) maximised

the cross-correlation of the stimulus with each of ten possible signals, or templates, by choosing the template that gave the highest correlation. The performance measure for this task, for both human and ideal observers, was the signal contrast energy necessary to obtain a specified performance criterion, in different levels of external noise, from low to high. The efficiency of the human observer, for face and texture discrimination tasks, was calculated as the ratio of human contrast thresholds to ideal observer contrast thresholds. Gold et al. found an increase in processing efficiency for face and texture discrimination over time, with no discernible changes in equivalent input noise. To trace the changes in observer calculations as a result of learning, Gold et al. additionally used the method of reverse correlation, described next.

1.2.4 Reverse correlation

Reverse correlation (also known as the classification image technique) originated in the auditory research of Ahumada (1967, cited by Ahumada, 2002) and has experienced a re-birth in recent years. Since its first application to vision research (Ahumada, 1996), the technique has been applied to a variety of perceptual tasks to estimate how observers use information in an image to reach a decision. The technique has subsequently been creatively applied to the perception of illusory and occluded contours (Gold et al., 2000), to face discrimination (Gold et al., 2004; Mangini and Biederman, 2004), vernier acuity (Beard and Ahumada, 1998) and spatial attention (Eckstein, Pham, and Shimozaki, 2004). Similar to the methods described previously, reverse correlation also presents a signal embedded in noise (see Figure 1.2), this noise is

used by the observer to reconstruct the information represented in memory for each classification. The outcome of a reverse correlation experiment is a classification image, an image that simply maps the correlation over many trials between the local noise contrast, at each image pixel, and the observer's response.

To provide an example of the technique in operation, consider a discrimination task where an observer has to discriminate between two signals, S_1 and S_2 . On each trial, the observer is presented with one of these signals embedded in luminance noise, and the contrast of the signal is adjusted to maintain a specified performance criterion. The task of the observers is to respond according to their perception of the presented stimulus. On some trials, the structure of the noise will lead the observer to incur errors; the noise may sometimes render the signal more similar to S_1 than to S_2 , or vice-versa. To determine the information that is represented by the observer to discriminate the signals, the technique classifies the noise patterns presented on each trial according to the four different stimulus response categories of SDT: hits, misses, false alarms and correct rejections, and for each category provides an average noise pattern. The linear combination of the noise samples that led to response hits and false alarms yields a *classification image*, clm_1 , which depicts the information that elicited those responses. A combination of the noise samples leading to response misses and correct rejections also yields a classification image, clm_2 , depicting the information that elicited those responses. The stimulus information an observer extracts to determine the presence or absence of a signal, or the features that enable discrimination between two signals, would be determined by the difference between clm_1 and clm_2 ($clm = clm_1 - clm_2$).

To establish the features of an object that are represented in memory, researchers in visual perception and cognition were forced to explore specific hypotheses regarding the presence or absence of an object property in memory by manipulating, sometimes crudely, one such feature (e.g. colour). An advantage of the reverse correlation technique is that it frees the experimenter from this constraint; a classification image, a pictorial representation of the information an observer extracts for a categorisation, is produced purely from the observer's classifications of the noise (plus signal) inputs. The technique has recently evolved to reveal internal representations of letters (Gosselin and Schyns, 2003) and stereo information (Gosselin, Bacon, and Mamassian, 2004) in experiments where only noise, and no signal, was presented. In these studies, observers were informed that a signal was embedded in the noise, and matched the noise-only input against a memorised template to reconstruct the signal they were asked to detect. The classification images from these experiments clearly depicted the image the observers believed they could detect in the noise, in the study of Gosselin and Schyns (2003), one of these images was a letter. Although the technique is not employed in the chapters to follow, its development is closely linked theoretically to the questions of interest and empirically in its methodological characteristics.



Figure 1.4. Examples of stimuli created from methodological techniques available to manipulate visual information. The left-most picture depicts a hybrid stimulus, that is a stimulus that comprises a neutral female face in high spatial frequencies and an angry male face in low spatial frequencies (adjusting viewing distance should alter observer perception between the two alternative percepts). The centre image depicts a sparse version of a female face, generated using the *Bubbles* technique. The right-most image depicts a face embedded in Gaussian noise; experimentally added noise is added to stimuli in experiments employing reverse correlation and noise-masking techniques.

Schyns, 2004). In the latter paper, Smith et al. (1994) applied the technique to determine the spatial scale information used to identify facial expressions. In

1.2.5 The Bubbles technique

The *Bubbles* technique of Gosselin and Schyns (2001) is a method designed to search for the information an observer uses to resolve a given categorisation task. On a trial-by-trial basis, the technique randomly samples an image generation space, e.g. a face, for information to present to an observer for classification. The image presented is, typically, a sparse version of a stimulus that is punctured by randomly located Gaussian-shaped holes¹; the middle picture of Figure 1.4 depicts such a stimulus. Over the course of an experiment the entire stimulus space is sampled. The observer's responses to these sparse stimuli are recorded according to the samples of information that lead to successful and unsuccessful, categorisations. The addition of these sparse samples together form images of the information leading to successful (hits and correct rejections), and unsuccessful

¹ The bubbles technique has been applied to stimuli spaces other than the 2-D image, usually using sampling methods other than Gaussian-shaped holes. Other such spaces have included phase information (Gosselin and Schyns, 2001).

(false alarms and misses) responses. Following a statistical analysis on these sets of information, a 'potent' image is constructed, that is the diagnostic subset of information used by the observer to resolve the task.

The technique has been applied to the categorisation of faces across the 2D image (Gosselin and Schyns, 2001) and across different spatial scales (Gosselin and Schyns, 2001; Schyns et al., 2002), to track the use of facial information through time (Vinette, Gosselin, and Schyns, 2004), to determine the facial information driving the ERP component, the N170 (Schyns, Jentzsch, Johnson, Schweinberger, and Gosselin, 2003), and in the information underlying the categorisation of emotional facial expressions (Smith, Cottrell, Gosselin, and Schyns, 2004). In the latter paper, Smith et al. (2004) applied the technique to determine the spatial scale information used to classify facial expressions. In examining this issue, Smith et al. used both human observers and a model theoretical observer to establish the efficiency with which observers used the information available. The human data revealed, across spatial scales, the facial information used to classify each of the six facial expressions, and the extent to which each expression correlated with the other expressions. A calculation of the human observer's 'optimal' use of information to resolve each expression categorisation was obtained, taking into account the diagnostic information of both the human and model observers. The technique has also recently been used to study information use in clinical populations such as a patient with damage to the amygdala that has resulted in an impairment to recognise fear from facial

¹ The bubbles technique has been applied to search spaces other than the 2D image, thereby using sampling methods other than Gaussian-shaped holes. Other search spaces have included phase information (Gosselin and Schyns, 2002).

expressions, explained by an inability to select information from the eyes (Adolphs, Gosselin, Buchanan, Tranel, Schyns, and Damasio, 2005).

As a tool to examine information use in perception and recognition, the bubbles technique has proved, in its young lifetime, a fruitful method in the toolbox of techniques. Few methods enable simultaneous sampling of image spaces such as the 2D image-plane together with spatial scales.

In Chapter 2, the technique is applied in a subjective task to determine the information underlying each stable percept of an ambiguous image.

1.2.6 Summary

The aim of the present section 'Tools of Information Use' was to provide an introduction to some of the techniques available to explore questions of information use in visual perception and recognition, and to those employed in the chapters to follow. Many of the techniques are complimentary, offering tools that promise a full account of the constraints and mechanisms in information use and its role in perception, attention and recognition. The suitability of these techniques to particular questions of information use ultimately depends on the question of interest. For example, the researcher interested in determining the information used across spatial scales for face recognition would be advised to consider the bubbles technique as opposed to the technique of reverse correlation since the former method is more suited to searching these two spaces simultaneously. On the other hand, if one is more interested in establishing the mechanisms involved in extracting information, at this point of development, the methods of noise masking and ideal observer analyses would be more

appropriate. An advantage of both the bubbles and reverse correlation methods, which in part seek to determine the stimulus information driving an observer's response, is that they can both be used in conjunction with ideal observer analyses, as is the method of noise masking. A disadvantage with both the bubbles and reverse correlation methods, however, relates to the large number of trials necessary to search an entire stimulus space, a consideration particularly relevant in studying specialist clinical populations. The application of these techniques, however, to questions relating to information use is still in its infancy, but the mass interest in their use hopefully will lead to a solution in reducing the number of trials such experiments demand.

1.3 Spatial Frequency Channels in Early Vision

1.3.1 Spatial frequency channels in the early visual system

An early stage in visual processing is the analyses of visual input by a bank of multiple channels² that independently process different spatial aspects of the input. In their seminal paper, Campbell and Robson (1967) applied a linear systems approach to establish the extent to which the contrast sensitivity function (CSF) could be used to predict contrast thresholds. The CSF simply describes an observer's sensitivity to a sine-wave grating as a function of its spatial frequency³. The results of measuring contrast detection (and

² A channel refers to a filtering mechanism that admits a restricted range of information (DeValois and DeValois, 1990); in the present context this information is spatial frequency information.

³ Spatial frequencies are expressed as the number of cycles per degree of visual angle and/or cycles per image. Whereas the former is a relative measure accounting for the distance from which an object is viewed, the latter is an absolute measure relating only to the information contained within the image.

discrimination) thresholds for gratings with various waveforms (sine-, square-, saw-tooth and rectangular-wave patterns) of identical spatial frequencies indicated that the contrast threshold for detecting the square-wave grating depended upon the contrast of the fundamental sine-wave component of the square-wave. The detection (or discrimination) of a visual input, therefore, is determined by the contrast of its components rather than its overall contrast. Campbell and Robson concluded that the visual system comprised multiple spatial channels, each sensitive to a limited range of spatial frequencies.

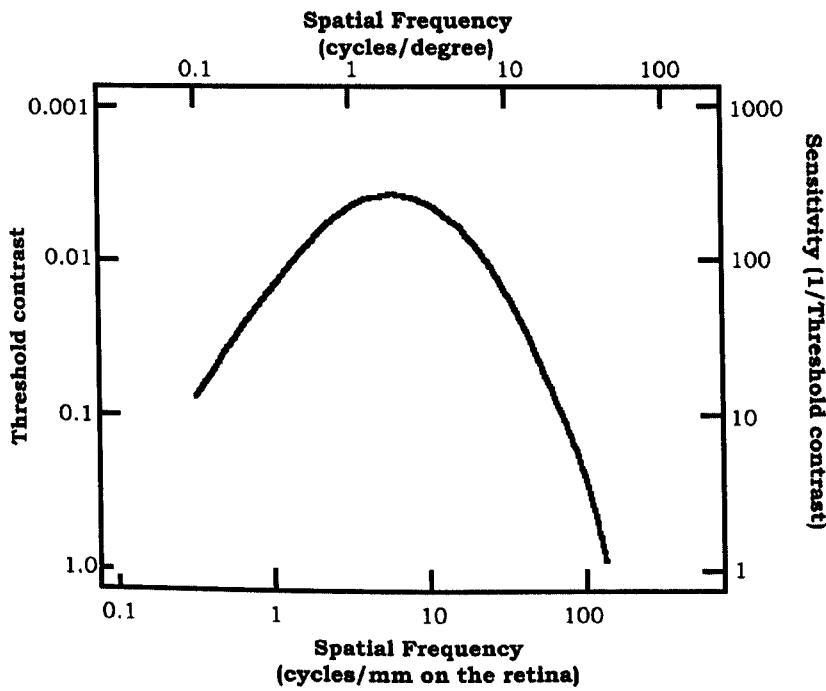


Figure 1.5. The Contrast Sensitivity Function (CSF) of an adult human observer. The threshold contrast required to detect a stimulus is plotted against the spatial frequency of the stimulus (figure adapted from Sekuler and Blake, 1994).

Blakemore and Campbell (1969a, b; see also Pantle and Sekuler, 1968) further substantiated this multiple channel theory in measuring the CSF before and after adaptation. The effect of adaptation, that is reduced sensitivity to a visual pattern following prolonged exposure, incurs temporary response suppression of the mechanisms responsible for their processing. Blakemore and Campbell found that adaptation to a particular spatial frequency increased contrast thresholds on the CSF only for a limited range of neighbouring frequencies of the adapting frequency, providing further support that the visual system analyses input with independent spatial frequency and orientation-selective channels.

As a result of these seminal research papers, the characteristics of these channels have been extensively studied. Research has revealed that the visual system comprises six such channels (Wilson, McFarlane, and Phillips, 1983), which are between one and two octaves wide⁴ (Wilson et al., 1983; for a comprehensive review of spatial vision see DeValois and DeValois, 1990 and for a historical perspective on spatial channels, see Wilson and Wilkinson, 1997). The channels are not entirely independent, as there are interactions (Henning, Hertz, and Broadbent, 1975) and some non-linearity (Henning, Hertz, and Broadbent, 1975; Stromeyer and Klein, 1975; Wilson and Bergen, 1979). Therefore, although there is dispute regarding the characteristics of these multiple spatial filters, there maintains a general acceptance that the visual system does process visual input via these orientation-selective spatial frequency bandwidth channels, and that these channels operate early in visual processing. The pioneering work of

⁴ An octave is a factor of two.

Campbell and Robson's subsequently compelled researchers to consider the spatial frequency composition of experimental stimuli.

1.3.2 Beyond Gratings: Spatial Frequency Channels and Recognition

The usefulness to the visual system of the existence of multiple spatial filters lays in the contribution of the output from these channels to the detection and discrimination of input more complex than gratings. How the visual system uses spatial frequency information in resolving complex tasks such as object, face and scene recognition has been the focus of extensive study in the past two decades. Most studies have focussed on determining the band(s) of spatial frequencies used by the visual system to resolve tasks such as letter identification (Gold et al., 1999a; Majaj, Pelli, Kurshan, and Palomares, 2002; Parish and Sperling, 1991; Solomon and Pelli, 1994), face recognition (Costen, Parker, and Craw, 1996; Fiorentini, Maffei, and Sandini, 1983; Gold et al., 1999a; Hayes, Morrone, and Burr, 1986; Schyns and Oliva, 1999) and scene recognition (Oliva and Schyns, 1997; for a review see Morrisson and Schyns, 2001).

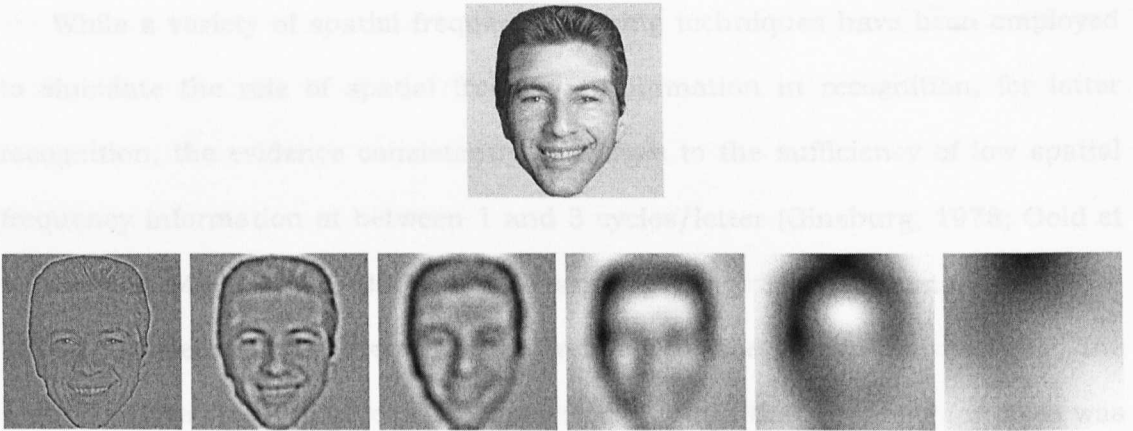


Figure 1.6. An example of spatial scales. The full-spectrum of spatial scales is shown in the top image. The bottom pictures show the result of a filtering process that decomposes the top image into six different spatial scales, one octave wide, from left-to-right, fine to coarse scales (in cycles per image, respectively, 128, 64, 32, 16, 8 and 4).

To illustrate the analyses of a complex image passing through multiple spatial channels, Figure 1.6 depicts an example of a complex image represented at multiple spatial scales. The original face (top picture) shows the full range of spatial information contained within the image. The six bottom pictures depict the outcome of applying a one-octave wide spatial filter to the image decomposing it into six different bands of spatial frequencies, high-to-low, from left-to-right. Note the information content contained within these different spatial frequency bands. Whereas finer facial detail such as wrinkles and eyebrow shape is available in the high spatial frequencies (left-most images), this information is absent in the lower spatial frequencies (right-most images), providing a more coarse representation of the face such as its shape. In exploring how spatial frequency information contributes to complex pattern recognition, many studies have focussed on determining the channel(s) involved in the identification of stimuli such as faces and letters.

While a variety of spatial frequency filtering techniques have been employed to elucidate the role of spatial frequency information in recognition, for letter recognition, the evidence consistently converges to the sufficiency of low spatial frequency information at between 1 and 3 cycles/letter (Ginsburg, 1978; Gold et al., 1999a; Majaj et al., 2002; Parish and Sperling, 1991; Solomon and Pelli, 1994). In determining the channel(s) related to face recognition however, the picture is less clear. Whereas Ginsburg (1980) found that accuracy for faces was best at a band of low spatial frequencies centred at 4 cycles/face width, Tieger and Ganz (1979) found greater masking effects at higher spatial frequencies (17 cycles/face width). Fiorentini et al. (1983) suggested that both low and high spatial frequencies were important for face recognition but that there were fewer errors for high-passed filtered images (containing frequencies higher than 5 cycles/face width). Later studies continued to find differences in face recognition performance; Rubin and Siegel (1984) reported optimal performance with low spatial frequency information at 1 cycle/face width, while others have proposed a central band of frequencies between 8 and 16 cycles/face width (Costen et al., 1996; Gold et al., 1999a; Parker and Costen, 1999), or between 20 and 25 cycles/face width (Hayes et al., 1986). These divergent results between the role of low and high spatial frequency information in face recognition are explicable in terms of the range of methodologies and the suitability of the identification tasks used (Morrisson and Schyns, 2001).

In establishing the spatial frequency channel(s) that contribute to resolving complex recognition tasks, questions have arisen as to whether this information is used in a stimulus-driven, bottom-up fashion, or rather is influenced by top-down processing. To determine the channel(s) mediating letter identification

Majaj et al. (2002) applied the technique of critical band masking (Solomon and Pelli, 1994) to letters of a variety of fonts, alphabets and sizes. In these studies, the method of critical band masking was used to determine the critical band of frequencies important for letter identification. The method involves measuring contrast thresholds for letter identification in the presence of low- or high-pass noise. Different predictions can be formed regarding the effects on contrast thresholds when observers are forced to use a single channel (bottom-up approach) versus the freedom to switch channels (top-down approach). If observers are constrained to use only a single critical channel, the sum of the contrast threshold elevations for letter identification caused by the low- and high-pass noise should equal the threshold increase affected by their sum⁵. In the case where observers are free to switch channels, thus improving the signal-to-noise ratio in a less noisy channel, low- and high-pass noise masking is expected to be less effective in elevating thresholds. The results supported the single channel model for letter recognition; threshold elevations suggested that observers were unable to switch channels to avoid noise. Rather, the channel used by observers to resolve the task of letter identification was determined “solely by the properties of the signal,” (p.1180). Letter identification, concluded Majaj et al. (2002), is bottom-up, “observers are not free to choose which channel they use” (p.1165).

Conversely, scene (Oliva and Schyns, 1997) and face categorisation (Schyns and Oliva, 1999) studies have suggested that the channel(s) observers attended to varied with categorisation task demands, purporting a top-down modulation of spatial-frequency processing that is in conflict with the conclusions of Majaj et al.

⁵ This effect is referred to as noise additivity (Pelli, 1981; cited in Majaj et al., 2002).

(2002). Using hybrid stimuli that combines two unique stimuli (an example of a hybrid stimulus used by Schyns and Oliva (1999) is shown in Figure 1.2a), for example, a neutral female face in high spatial frequencies (HSF) and an angry male face in low spatial frequencies (LSF), Oliva and Schyns (1997; see also Schyns and Oliva, 1999) sought to demonstrate that the information demands of a categorisation task can influence the spatial scale to which an observer attends. Oliva and Schyns generated hybrid stimuli that comprised a scene represented in low-spatial frequencies (e.g. a highway) and another scene represented in high-spatial frequencies (e.g. a city). Prior to the testing phase in which observers were presented with these hybrids for categorisation, observers were allocated to a HSF group or a LSF group where they were presented with hybrid stimuli that combined scene information at one scale, e.g. a highway in HSF, with noise in the other scale, e.g. LSF noise. This sensitisation phase required observers to attend to the scale containing meaningful scene information. Categorisations of hybrid stimuli in the subsequent testing phase revealed that participants' perception of the hybrid stimuli aligned with the scale information to which they were sensitised in the previous stage. Observers who were sensitised to hybrids comprising a LSF scene and HSF noise, on presentation of a hybrid with scene information at both scales, reported a perception of the hybrid consistent with their sensitisation, that is the scene represented in LSF. The information demands of the categorisation task in the sensitisation phase, the diagnostic scale information being present at only one scale, biased perception of the stimulus in the testing phase. Scale selection, according to Oliva and Schyns, was therefore determined by the presence of diagnostic information, rather than

being fixed at a single scale. This conclusion conflicts with that of Majaj et al. (2002) who concluded that scale selection was not flexible, but fixed.

In considering the findings of Majaj et al. (2002), Sowden et al. (2003) suggest that the failure to find evidence of the ability to select different channels may be due to the task and stimuli used by Majaj et al. Letter identification, which generally takes place in high contrast, low external noise conditions, is a highly practised task for which a single channel may be sufficient; channel switching to avoid noise may therefore require substantial unlearning. While the observers of Majaj et al. were not explicitly directed to attend to a specified spatial frequency band, the observers in a study by Sowden et al. (2003) received information regarding the channel to which they should attend. Sowden et al. conducted a study investigating the top-down control of spatial frequency processing using auditory pre-cueing to manipulate attention to particular spatial frequency channels. Observers were presented with plaid stimuli comprising a low spatial frequency grating at one orientation and a high spatial frequency grating at another orientation. The task of the observer was to judge the orientation of the spatial frequency grating. The auditory tone, presented prior to the onset of the plaid stimuli, informed observers of the spatial frequency of the upcoming grating. It was predicted that attending to the channel *expected* to carry the task-relevant information would result in the observer reporting the orientation of the cued grating. The results of this study showed that the perception of plaid stimuli was influenced by the auditory pre-cue, indicating that observers experienced a selective perception of the plaid stimuli that was consistent with the cue; observers were able to monitor the channel they expected would carry task-relevant information. In contrast to Majaj et al. (2002), Sowden et al. (2003)

conclude top-down attentional modulation of spatial frequency channels, where observers *are* free to choose the channel that carries task-relevant information. This conclusion is further supported by studies showing a differential use of spatial frequency information as a function of task demands using the bubbles technique (Gosselin and Schyns, 2001; Schyns et al., 2002) and the N170 component (a negative peak around 170 ms post-stimulus) of the visual event-related brain potential (ERP; Goffaux, Jemel, Jacques, Rossion, and Schyns, 2003). These studies showed that observers attended to different spatial frequency channels to extract the diagnostic information relevant to resolve the task.

In Chapter 2, I apply the bubbles method of Gosselin and Schyns (2001) to the 2D image plane of a bi-stable image, and the third dimension of spatial scale, to determine, within and across each spatial scale, the information underlying each percept. Consistent with the perspective of Sowden et al., my position is that observers will monitor different spatial frequency channels if task-relevant information is present. Additionally, in Chapter 3, as a methodological tool to explore questions of information use, I use hybrid stimuli. Such stimuli are useful in addressing questions of information use because by combining multiple stimuli into a single stimulus, enables the stimulus itself to become its own control. Such stimulus manipulations have contributed to research on the effects of task demands in information selection in behavioural (Oliva and Schyns, 1997; Schyns and Oliva, 1999) and electrophysiological studies (Goffaux, Jemel, Jacques, Rossion, and Schyns, 2003). Other types of hybrid stimuli include chimeric face stimuli, that is a stimulus comprising a face of one gender on the left side and that of another gender in the right side. Burt and Perrett (1997; see

also Butler, Gilchrist, Burt, Perrett, Jones, and Harvey, 2005) generated such stimuli to test perceptual asymmetries of different face categorisations and found that participants used information from the left side of faces to resolve categorisations of gender, age, attractiveness, speech and expression.

In relation to the flexibility of the observer to use the information available for perception and recognition, an additional relevant factor is the consideration that there may be a modulation in information use as a result of learning. For example, an observer may use information x when the signal is largely unfamiliar, but over time may change to using information y . The next section reviews research on the phenomena of perceptual learning, and the mechanisms that have been proposed to enable learning. At the heart of this research are the questions: What is perceptual learning? What are the mechanisms enabling learning? These questions are closely related to attention to information for perception and recognition because to extract diagnostic task-relevant information to resolve a task, the observer must firstly learn what this information is in order to extract it.

1.4 Perceptual learning

In his chapter *Discrimination and Comparison*, James (1890) recounts real world examples of perceptual learning; described is the ability of a man to “distinguish by taste the upper and the lower half of a bottle of old Madeira,” and “another will recognize by feeling the flour in a barrel, whether the wheat was grown in Iowa or Tennessee,” (p.509). At his time of writing, James suggested that psychologists were so familiar with the phenomenon of perceptual learning that they didn’t feel the need to seek to explain it. Indeed, he states, “At most

they have said: "Attention accounts for it; we attend more to habitual things, and what we attend to we perceive more minutely," (p.510). James's response to this dismissal of accounting for the phenomenon of perceptual learning was a call for a "less general" explanation in favour of one "more precise". In recent years, strong attempts have been made to characterise precisely the conditions necessary for perceptual learning to occur, its permanence, and its underlying mechanisms; a review of this research follows.

1.4.1 The Nature of Perceptual Learning

In considering the nature of perceptual learning, here we refer only to those studies that are, in essence, purely perceptual. For example, while the categorisation tasks of Gauthier, Williams, Tanaka, and Tarr (1998) undeniably entail a perceptual component in learning to discriminate computer generated 'greebles', their tasks also required considerable semantic memory involvement for the names of individual greebles, their gender and family names, and may therefore encompass additional processes; such studies are omitted here to exclusively focus on studies isolating perceptual learning. Empirically, perceptual learning has been shown to occur for a variety of visual discrimination tasks including the discrimination of vernier acuity (Fahle and Edelman, 1993; Poggio, Fahle, and Edelman, 1992), an object's orientation (Doshier and Lu, 1998, 1999; Lu and Doshier, 2004) and motion trajectory (Ball and Sekuler, 1987; Fahle and Morgan, 1996; Liu and Weinshall, 2000), stereoacuity (Sowden, Davies, Rose, and Kaye, 1996), gratings (Fine and Jacobs, 2000; Fiorentini and Berardi, 1980), texture (Gold et al., 1999b, 2004; Karni and Sagi, 1991), face (Gold et al., 1999b,

2004), letter (Tjan, Chung, and Levi, 2002a, b), and position discrimination (Li, Levi, and Klein, 2004). The learning is often found to be specific to training conditions, failing to transfer to other stimuli, tasks, visual field position, and eye of training (Ahissar and Hochstein, 1993, 1997; Fahle, Edelman, and Poggio, 1995; Fahle and Morgan, 1996; Fiorentini and Berardi, 1980; Karni and Sagi, 1993; Poggio et al., 1992; although see Liu and Weinshall, 2000 for evidence of generalization of learning in a motion discrimination task). Perceptual learning can occur following relatively short training periods (Fahle et al., 1995; Fiorentini and Berardi, 1980; Poggio et al., 1992), to requiring longer exposure (Karni and Sagi, 1993). Perceptual learning effects have been obtained in the absence of feedback (Ball and Sekuler, 1987; Fahle et al., 1995), and perhaps surprisingly, to task irrelevant features exclusive of an observer's explicit awareness and attention (Watanabe, Náñez, and Sasaki, 2001; Watanabe, Náñez, Koyama, Mukai, Liederman, and Sasaki, 2002).

Psychophysical studies indicating the specificity of perceptual learning to restricted stimuli, tasks and retinal locations, imply that the neural modulations underlying perceptual learning may occur in early stages of the visual system (Fahle, 2004; Karni and Sagi, 1991). Physiological evidence also suggests that plasticity in the adult primary visual cortex may support the neuronal changes underlying perceptual learning (see Fahle, 2004 for the case supporting modulations of early cortical mechanisms). Indeed, Watanabe et al. (2002) found greater plasticity in lower-level motion processing brain areas than higher-level motion brain areas in a passive perceptual learning task. At a behavioural level, functional descriptions of the changes that take place in these mechanisms with

perceptual learning have been investigated within a signal detection theory framework; an overview of this research follows.

1.4.2 Potential mechanisms of perceptual learning

To characterise the mechanisms responsible for perceptual learning, it has proved useful to consider the problem within the framework of SDT (Doshier and Lu, 1998, 1999; Gold et al., 1999b, 2004; Lu and Doshier, 2004; Tjan, Levi, and Chung, 2002). In appealing to SDT to elucidate the effects of perceptual learning, independent measures of an observer's visual sensitivity and decision-making processes to a stimulus event can be traced through time. How does perceptual learning alter an observer's sensitivity? The SDT framework enables the partitioning of an observer's visual sensitivity into measurable and unique mechanisms. The first mechanism that may be altered with perceptual learning is the efficiency with which an observer samples the visual information available termed *signal (or stimulus) enhancement*. How this stimulus enhancement mechanism achieves this perceptual learning effect varies according to the perspectives of different researchers (Gold et al. 1999b, 2004; Doshier and Lu, 1998, 1999) and will be discussed in subsequent sections. A second mechanism enabling perceptual learning, termed *internal noise reduction*, may operate to alter the strength of the internal noise naturally associated with sensory processing. This mechanism functions to decrease internal noise, and would also result in improved sensitivity. A third possible mechanism that enables perceptual learning, *external noise reduction*, has also been invoked (Doshier and Lu, 1998,

1999). While deemed unnecessary by some researchers (Gold et al., 2004), such a mechanism would operate to exclude any external noise in the stimulus.

To distinguish between these potential mechanisms the method of noise masking has been applied to the perceptual learning tasks of orientation (Doshier and Lu, 1998, 1999), texture (Gold et al., 1999b, 2004), face (Gold et al., 1999b, 2004), and letter (Tjan et al., 2002) discriminations. For each potential mechanism, the noise masking method provides unique performance signatures when plotting an observer's contrast threshold, at a specified performance criterion, for a signal embedded in a range of noise densities (Doshier and Lu, 1998, 1999; Gold et al. 1999b, 2004). Figure 1.7 depicts hypothetical noise-masking functions where contrast energy thresholds are plotted as a function of the density of external noise; note how the contrast energy threshold increases as a function of increases in external noise density. The figure depicts the changes that would occur in the noise masking function for a perceptual learning mechanism that in a) augments the strength of the incoming signal, b) diminishes the magnitude of internal noise, and c) improves performance by excluding external noise.

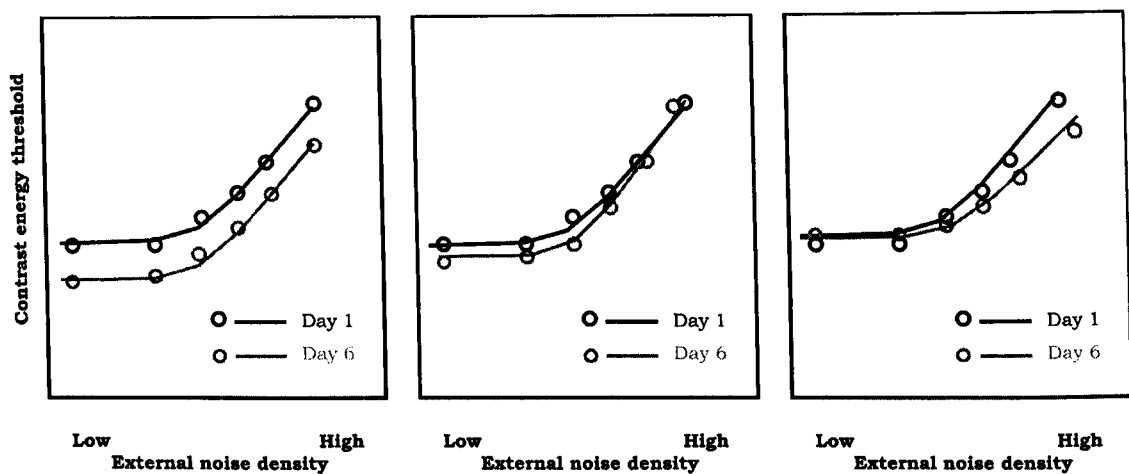


Figure 1.7. Hypothetical noise masking functions for three potential mechanisms of perceptual learning. Each panel indicates the potential effect of perceptual learning on the noise masking function characterised by a mechanism of signal, or stimulus enhancement in the left panel, internal noise reduction in the central panel, or external noise reduction in the right panel.

These mechanisms have, over the years, been articulated in various forms by researchers in different fields, including physiology and engineering. Although similar methodology is employed between researchers, to characterise the changes that occur with perceptual learning, the mechanisms proposed vary depending on the theoretical leanings of different research groups. There is agreement that sampling efficiency and internal noise reduction are two potential candidates. However, there exists controversy pertaining to the early versus late addition of noise into the observer's system, and the contention that two as opposed to three mechanisms are sufficient to account for perceptual learning. Note that different researchers use similar terminology for these mechanisms, while the changes these mechanisms effect on the noise masking function vary.

In the perceptual learning literature, the early noise model, propounded by Gold et al. (1999b, 2004), posits two distinct processes to explain perceptual

learning: signal enhancement (otherwise referred to as *sampling efficiency* or *calculation efficiency*) and internal noise reduction. Doshier and Lu defend an alternative position, a late noise model, suggesting a full explanatory account of perceptual learning requires a third mechanism, external noise reduction, in addition to mechanisms of signal enhancement and internal noise reduction. While the methodology, and experimental results, of Gold et al. and Doshier and Lu are highly similar, the subsequent interpretations and models constructed to account for these results differ. An introduction to the models and arguments of Gold et al. and Doshier and Lu, as examples of early and late noise models, respectively, follow in the next section. A brief functional description of each of the proposed mechanisms and how they are manifested on noise-masking functions are described.

1.4.2.1 An early noise model: Gold et al. (1999b, 2004)

In characterising the mechanisms enabling perceptual learning, Gold et al. appeal to the black-box observer model of Pelli (1981), the Linear Amplifier Model (LAM), depicted in Figure 1.8. In resolving a perceptual task, this simple observer model depicts the observer receiving a noisy stimulus as input. To the representation of this noisy input the observer contributes another noise source, a fixed amount of *contrast-invariant internal noise*. Following a *contrast-invariant calculation* on the stimulus representation, the observer finally reaches a perceptual decision regarding the stimulus input that is based on the internal response. To note in the model is that the internal noise source is added early in processing, prior to the calculation stage.

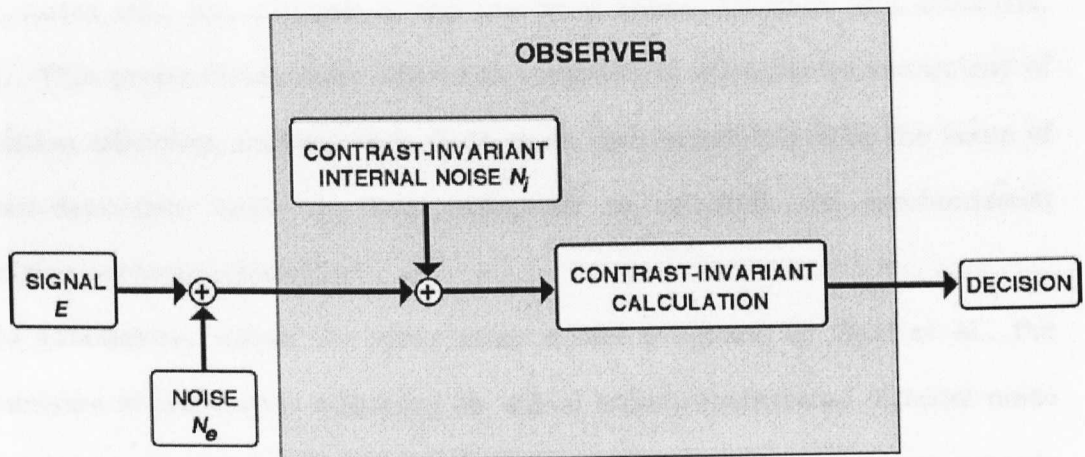


Figure 1.8. Pelli's black-box model of a human observer in a perceptual task. The observer in this model receives as input a signal embedded in noise, to which a fixed amount of internal noise is added. The observer performs a calculation producing an internal response, the strength of which determines the resulting perceptual decision (an adapted version of Pelli's model, from Gold et al., 2004).

Within this model, perceptual learning potentially affects *contrast-invariant internal noise*, via a mechanism of internal noise reduction that operates to adjust an internal noise source that is invariable to the strength of the visual input. Alternatively, or in addition, the *contrast-invariant calculation*, also indifferent to the strength of the incoming signal, may be affected by a mechanism of signal enhancement. This calculation stage denotes the internal signal strength, which on comparison with an ideal observer is a measure of the proportion of available information used by the human observer.

Gold et al. also acknowledge an additional, and separate, internal noise source, *contrast-dependent noise*, excluded from Pelli's model (Pelli, 1990 asserts

that this noise source results from variability in the observer's calculation). The proportion of this contrast-dependent noise, unlike contrast-invariant internal noise, varies with the strength of the incoming signal (Burgess and Colborne, 1988). This proportional noise source is confounded with the measurement of calculation efficiency, consequently Gold et al. deal separately with the issue of contrast-dependent noise in their endeavour to establish the mechanism(s) underlying perceptual learning.

To summarise, within the early noise model proposed by Gold et al., the mechanisms of calculation efficiency (or signal enhancement) and internal noise reduction are the potential candidates enabling the performance improvements associated with perceptual learning. To distinguish between these mechanisms, Gold et al. employ the techniques of noise masking and response consistency.

The method of noise masking is used to isolate any change in calculation efficiency, that is, how efficiently the observer uses the information available, as a result of perceptual learning. This method enables the calculation of the contrast threshold of the human observer, and the theoretical ideal observer, in varying levels of noise density. Whereas the contrast thresholds for both the human and ideal observers will vary according to the discriminability of the stimulus set (and will increase as the density of external noise increases), only the human observers contrast threshold will be affected by the magnitude of internal noise and their efficiency in using the information available. The measure of *calculation efficiency*, that is the proportion of information used by the human observer, is obtained by comparing the contrast threshold of the human observer to that of the ideal observer. According to the early noise model of Gold et al. a mechanism that functions to *increase calculation efficiency* will reduce contrast thresholds

uniformly across all external noise levels. Whereas, a mechanism that operates to *reduce internal noise* will reduce contrast thresholds only at low levels of external noise, causing a downward shift in the kink-point of the function (the left and centre panels of Figure 1.7 depict the changes in the noise-masking functions associated with each of these potential perceptual learning mechanisms).

To establish the mechanisms underlying the perceptual learning of face and texture pattern discrimination, Gold et al. employed the methods of noise masking, response consistency and reverse correlation. Firstly, in their noise making experiment, observers learned to discriminate, over six learning sessions, ten faces and four texture patterns of varying contrast in different levels of experimenter added external noise. An ideal observer employing an optimal decision rule cross-correlates the noisy input signal with each of the noise-free template stored in memory, and responds according to the signal yielding the highest cross-correlation. Their results from tasks of face and texture discrimination, depicted in Figure 1.9, show increases in contrast thresholds as a function of external noise, and a decrease in contrast thresholds across all external noise levels, as a result of practice for both face and texture discrimination tasks. Moreover, this effect of learning on the noise masking functions revealed increases in calculation efficiency for both tasks; an increase in efficiency of a factor of 4 in discriminating faces, and by a factor of 2 to 3 in texture discrimination. Since a mechanism operating to reduce contrast-invariant internal noise, within this model, produces a downward shift only in low levels of external noise, Gold et al. rule out a contribution of this mechanism to the changes occurring with perceptual learning. However, recall that the Linear

Amplifier Model (Pelli, 1990) does not enable changes in contrast-dependent internal noise to be distinguished from changes in calculation efficiency.

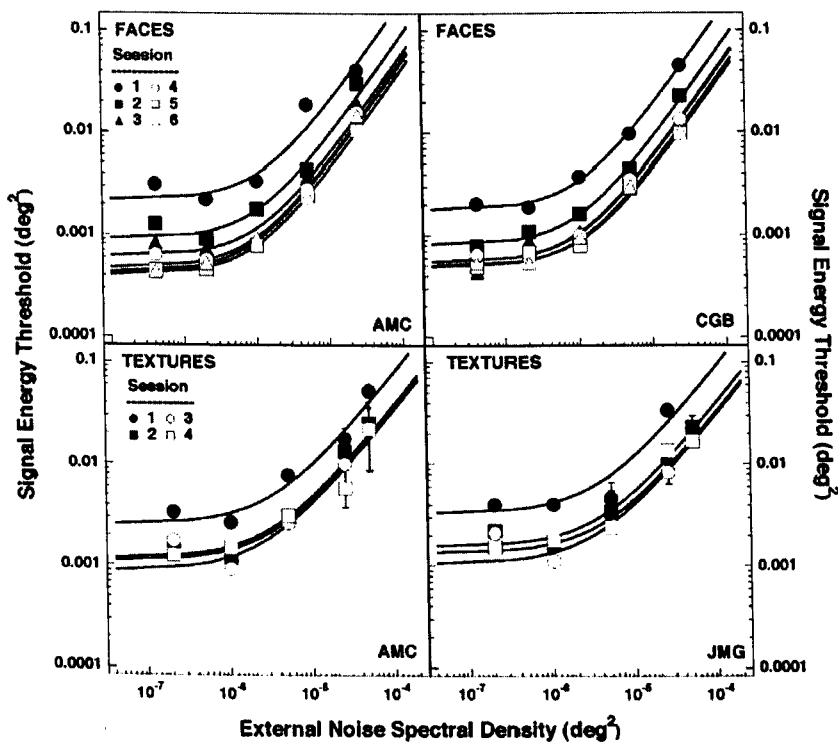


Figure 1.9. Noise masking functions from Gold et al. (1999b, 2004). The upper panels depict noise-masking functions of two observers for the task of face discrimination, and the lower panels, the task of texture discrimination. The change in performance for each learning session, downward shifts in contrast energy thresholds across all levels of external noise signify a mechanism operating to enhance the strength of the incoming signal.

To exclude the possibility of changes in contrast-dependent noise contributing to the changes observed in calculation efficiency with learning, Gold et al. employ a technique additional to noise masking, the technique of *response*

consistency, in a separate experiment. This technique, similar to the method of noise masking, measures contrast energy thresholds for stimuli embedded in external noise. The difference however is that this technique involves presenting twice an identical sequence of trials comprising stimuli embedded in identical noise patterns of low noise density (where contrast-invariant noise dominates) or high noise density (where contrast-dependent noise dominates) and measuring the consistency in responses between the two passes through an identical experiment. For an ideal observer, the responses between these two passes will be identical. However, for a human observer with internal noise, there will be some inconsistency between these responses to identical stimuli. The degree of inconsistency between the two passes is determined by the internal/external noise ratio. The contribution of contrast-invariant noise to any response inconsistencies is eliminated using high levels of external noise where this noise source has little effect, therefore isolating the contrast-dependent noise source. This response consistency technique subsequently permits the distinction between a perceptual learning mechanism that reduces contrast-invariant noise and one that improves calculation efficiency by reducing contrast-dependent noise. The results from this experiment showed that response consistency did not change at high levels of external noise, indicating that contrast-dependent noise did not significantly change as a function of perceptual learning.

Taken together, the noise masking and response consistency results from Gold et al. revealed no changes in observers' contrast-invariant and contrast-dependent internal noise. Rather, the results suggested that the mechanism underlying the perceptual learning of faces and texture patterns is a mechanism of signal enhancement (or calculation efficiency) that operates to increase the

observer's efficiency in extracting the information available in the stimulus to resolve the task.

1.4.2.2 A late noise model: Doshier and Lu (1998, 2000)

Doshier and Lu (1998, 1999) developed the Perceptual Template Model (PTM) to characterise the mechanisms underlying the performance improvements observed in perceptual learning tasks, and in tasks of selective spatial attention (Lu and Doshier, 1998, 2000, 2004). The PTM, depicted in Figure 1.10, extends the LAM with its inclusion of *transducer non-linearity* and *multiplicative internal noise* (or contrast-dependent internal noise). In addition, whereas the PTM can account for performance at different criterion levels, the LAM cannot.

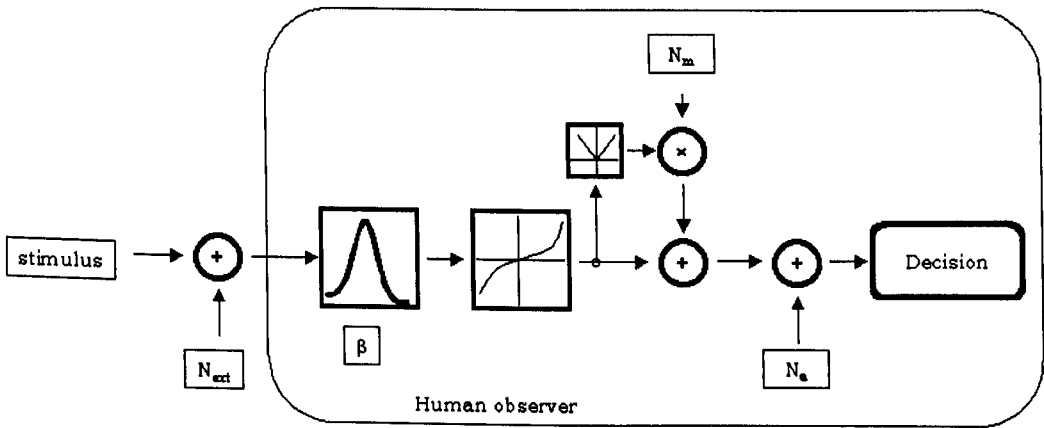


Figure 1.10. The Perceptual Template Model (PTM) of Doshier and Lu. The observer model consists of four components: (1) a perceptual template (β); (2) a multiplicative internal noise source (N_m); (3) an additive internal noise source (N_a); and (4) a decision process.

An example of a late noise model, the PTM of Doshier and Lu invokes three potential mechanisms of perceptual learning, *stimulus (or signal) enhancement*, *external noise exclusion* and *internal noise suppression*, each uniquely affecting the noise masking function. The first mechanism, stimulus enhancement, involves augmenting the strength of an incoming sensory signal by a) increasing the gain on the output of the relevant perceptual template, or by b) reducing *additive internal noise* (termed contrast-invariant internal noise by Gold et al. 2004). A stimulus enhancement mechanism would effect change on the noise-masking function at low levels of external noise; remember additive internal noise has no contribution at high external noise levels, so increasing the gain of the perceptual template at high levels of external noise would also entail a simultaneous increase in the gain of the external noise. Notice that this mechanism may effect change in either or both a) increasing the gain on the output of the relevant perceptual template and b) reducing additive internal noise. The role of a mechanism operating to suppress multiplicative internal noise (or in the terminology of Gold et al., contrast-dependent noise) would be to diminish the effects of this noise source by improving performance in both low and high levels of external noise. Lastly, a PTM mechanism that functions to restrict the effect of external noise in the stimulus would effectively narrow or tune the filter or template processing the incoming signal to exclude competing external noise. The effect of this mechanism on the noise masking function would be to improve performance, reduce thresholds, in high external noise conditions. In narrowing the template to exclude external noise, the gain of the template is decreased (the converse effect of the stimulus enhancement mechanism).

A crucial difference between the models of Gold et al. and Doshier and Lu is the stage at which the observer's internal noise is added; for Gold et al. the observer's internal noise is added early in the model, prior to the calculation, whereas for Doshier and Lu the addition of noise occurs later, after the calculation. Also, for Gold et al. the mechanism of signal enhancement effects change only on the efficiency of the observer to use the information available, and is expressed on the noise masking function as a downward shift across both low and high noise levels

In their studies of perceptual learning, Doshier and Lu measured contrast thresholds for oriented stimuli embedded in varying levels of noise density for an orientation discrimination task presented in the peripheral visual field. Similar to Gold et al., the results of Doshier and Lu, depicted in Figure 1.11, showed that contrast thresholds increased as a function of increases in external noise density and decreased as a function of practice. The effect of learning on the noise-masking function for orientation discrimination produced a pattern of results similar to Gold et al., revealing a downward shift in contrast threshold across all densities of external noise. A difference in the observer models invoked by Gold et al. and Doshier and Lu in explaining perceptual learning have led to a difference in the interpretation of these similar results.

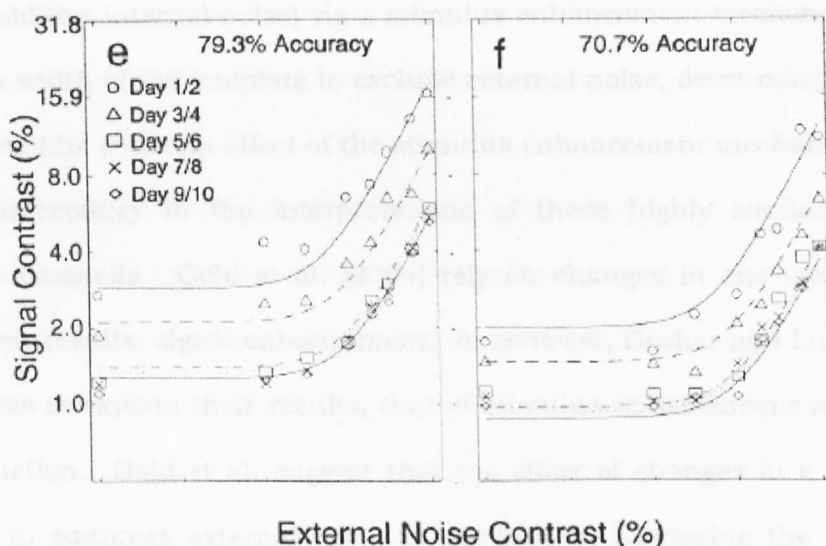


Figure 1.11. Noise masking functions from Lu and Doshier (1998). Each function represents experimental data on the mechanism(s) underlying performance improvements in spatial attention for two different response criteria. Here, the performance signature, a downward shift in the noise-masking function at all levels of external noise, according to Doshier and Lu corresponds to mechanisms of additive internal noise reduction and external noise reduction.

Recall that Gold et al. interpreted a similar pattern of results in face and texture discrimination tasks as evidence for a mechanism of signal enhancement, that is, an increase in calculation efficiency. Doshier and Lu, however, interpreted their results as evidence for a combination of mechanisms. The performance improvements at low levels of external noise, they suggest, were enabled by a stimulus enhancement mechanism operating to suppress additive internal noise, whereas the improvements observed at high levels of external noise were attributed to a mechanism of external noise exclusion. In other words, operationally the changes effected with perceptual learning in the studies of Doshier and Lu were to enhance the strength of the incoming stimulus by increasing the gain of the stimulus-relevant template (formally equivalent to

reducing additive internal noise) via a stimulus enhancement mechanism, and to narrow the width of the template to exclude external noise, decreasing the gain of the template (the converse effect of the stimulus enhancement mechanism).

The discrepancy in the interpretations of these highly similar results is difficult to reconcile. Gold et al. (2004) rely on changes in one mechanism to explain their results, signal enhancement. In contrast, Doshier and Lu invoke two mechanisms to explain their results, that of stimulus enhancement and external noise reduction. Gold et al. suggest that the effect of changes in a mechanism operating to suppress external noise is two-fold: in narrowing the template to exclude external noise, the effect is to make the template more efficient (similar in operation to the change effected with a signal enhancement mechanism), and reduces the gain of the output of the template (the converse effect of the stimulus enhancement mechanism). Gold et al. argue that the external noise exclusion mechanism of Doshier and Lu is really a combination of mechanisms, one whose role is to reduce the gain of the template, the opposite of the role of the stimulus enhancement mechanism also invoked by Doshier and Lu to account for their results. Gold et al. therefore maintain that their LAM-based account is more parsimonious in invoking a single mechanism, signal enhancement or calculation efficiency, to explain their perceptual learning results than the appeal to two mechanisms in the PTM of Doshier and Lu.

However, Lu and Doshier (2004) showed that perceptual learning improved performance in a foveal orientation discrimination task only in high levels of external noise. From this study, Lu and Doshier concluded that a mechanism of external noise exclusion enabled performance improvements in their perceptual learning task. In relation to the LAM model proposed by Gold et al. (2004), Lu

and Doshier point out that this model could only account for such data in terms of an increase in calculation efficiency, and as an *increase* in contrast-invariant internal noise.

In response to Lu and Doshier, Gold et al. acknowledge that, in this case, it is more parsimonious to assume that internal noise would remain unchanged, rather than increasing as a result of learning. Gold et al. suggest in this case that the noise should occur after the calculation, as it does in the PTM, rather than before.

In an attempt to explain the difference between their findings and those of Lu and Doshier, Gold et al. suggest that the difference may lie in the processing stages involved in the different tasks and the subsequent contribution of internal noise. Whereas, the gratings used by Doshier and Lu may reflect earlier processing stages, the face and texture discrimination tasks of Gold et al. may be processed by later stages in the visual system. In their letter discrimination task, Tjan et al. (2002a) accounted for performance improvements in terms of changes in internal (additive or contrast-invariant) noise reduction and external noise exclusion. In assessing both the LAM and the PTM in accounting for their data, Tjan et al. suggest that to account for criterion-independent performance, the transducer nonlinearity component of the PTM is necessary. Ultimately, the success of these models will depend on how well they account for data in different visual tasks. The future of this type of research intimates to be particularly fruitful, and surely will emerge to contribute to greater understanding of the mechanisms involved in enabling perceptual learning.

In Chapter 4, I apply the noise masking paradigm to the perceptual learning of face discrimination. In an attempt to trace over time the features or regions

being learnt to discriminate the faces, we measured contrast thresholds on each trial for each of six regions of the face embedded in varying levels of noise density.

1.4.3 Summary

This aim of the present section on Perceptual learning was to introduce the characteristics of perceptual learning, and the mechanisms that enable it. As a new area of research, the endeavour to functionally describe the behavioural mechanisms underlying perceptual learning is proving fruitful. While different observer models offer contrasting accounts of the data, the research in applying the technique of noise masking to determine the mechanisms of perceptual learning is still in its infancy (the first paper was published by Doshier and Lu, 1998). Already this research has focused on the perceptual learning of orientation discrimination (Doshier and Lu, 1998, 1999; Lu and Doshier, 2004), discrimination of faces and textures (Gold et al., 1999b, 2004), letters (Tjan et al., 2002), and spatial position (Li et al., 2004). The use of complementary techniques such as reverse correlation and bubbles to determine the regions of the signal that influence the observer's performance, in conjunction with noise masking, provide a fuller, more precise account of perceptual learning than previously possible. The potential convergence of these methods with electrophysiological (Liebe, Gold, Busey, and O'Donnell, 2004) and neuroimaging techniques (Tjan, Lestou, Bülthoff, and Kourtzi, 2004) make this a promising area of research.

Thus far, I have reviewed issues related to the information observers select to resolve different perceptual and recognition tasks. In the next section I review

research that examines the performance benefits of selectively attending to visual information. In addition to the effects of attention on behavioural performance, I also consider electrophysiological studies that inform us of the effects of attention on the time course of information processing.

1.5 Selective spatial attention

The spatial resolution of selective spatial attention, that is the selection of information across the visual field, has been extensively studied for decades. The shape of attention across the visual field has been previously described as being shaped like a 'spotlight' (Posner, Snyder, and Davidson, 1980), a 'zoom-lens' (Eriksen and St James, 1986), and a 'donut' (Müller and Hübner, 2002). The mapping out of the spread of attention across the visual field has important implications in constraining theories of selective attention to information for further visual processing. However, while these metaphorical accounts are constructed from robust patterns observed in experimental data, the employment of metaphors to describe attention has contributed to the notion that attention is a fuzzy concept. Indeed, the very word 'attention' as connoting imprecise processes has led some to under-use the term (Pashler, 1998). Yet, much progress has been made in understanding and characterising the spatial and temporal resolution of visual selective attention, and the functional nature of its underlying mechanisms (Baldassi, Burr, Carrasco, Eckstein, and Verghese, 2004).

Here, we review the findings of visual attention research, beginning with the experimental paradigm often used by researchers exploring the effects of spatial

attention across the visual field. Next, we consider the temporal resolution of visual attention explored in studies employing event-related brain potentials (ERPs), and review research that suggests that attention not only improves the discriminability of visual input but that attention also speeds up the rate at which visual input is processed.

1.5.1 Manipulating spatial attention: peripheral and central cueing

Several studies have consistently demonstrated performance benefits, reflected in increased accuracy, response time improvements, and enhanced perceptual sensitivity, when observers are given advance information regarding target location (Bashinski and Bacharach, 1980; Downing, 1988; Henderson, 1996; Posner, Snyder, and Davidson, 1980). In such studies, a typical experimental set-up would consist of a pre-cue directing the attention of the observer to a particular spatial location. The information provided by the pre-cue is manipulated to compare performance across conditions where the cue provides a) accurate information regarding the target in valid cueing conditions, b) inaccurate information regarding the target with invalid cues, and c) no information regarding the target in neutral cueing conditions. Such studies show greatest facilitation effects for valid cues, least facilitation effects with invalid cues, and intermediate effects for neutral cues (Müller and Rabbitt, 1989). The probability of the cue indicating target location is determined by the experimenter, and is set, conventionally, between $p(.5)$ to $p(.75)$. The presentation of the target stimulus typically follows the pre-cue after a variable time interval

(referred to as cue-target stimulus onset asynchrony, SOAs) ranging from short temporal delays to longer delays.

The pre-cueing procedures used to manipulate attention have been broadly classified into two types: central and peripheral. Research has demonstrated that these cue types invoke two distinct attentional systems (e.g. Cheal and Lyon, 1991; Müller and Rabbitt, 1989; Posner, 1980). The peripheral cue, typically presented adjacent to the peripheral target location, engages an exogenous mechanism and the latter, presented centrally at fixation and pointing to the target location, activates an endogenous system (Posner, 1980). There is evidence to suggest that these two systems exhibit functional differences and yield different performance benefits. Functionally, the exogenous system is said to be reflexive, engaged by brusque visual onsets appearing in the visual field, and is characterised by a transient response. Conversely, the endogenous system is considered voluntary, operating consciously according to goals, and is characterised by a sustained response profile.

In relation to facilitating performance, the exogenous system produces relatively greater effect sizes than the endogenous system at short temporal intervals between the offset of the cue and the onset of the target stimulus; with peripheral cues peak facilitation is observed at SOAs of between 100 and 175 ms, compared to between 400 and 725 ms with central cues (Henderson, 1991; Müller and Rabbitt, 1989). Inhibitory effects of cueing are also evident in peripheral cueing conditions while being absent in central cueing conditions (Cheal and Lyon, 1991; Müller and Rabbitt, 1989; Posner and Cohen, 1984; Rafal, Calabresi, Brennan, and Sciolto, 1989); studies engaging the exogenous system demonstrate inhibition of return (IOR), that is, an inhibitory effect

exhibited in slower reaction times for targets at previously cued locations than for targets at novel locations. The effect typically becomes evident at 300 ms after the presentation of a peripheral cue and lasts as long as 2000 ms (Posner and Cohen, 1984). This inhibitory effect under peripheral cueing conditions is unaffected by the probability of the cue, and has recently been shown to affect not only the previously cued location but beyond, encompassing the cued hemifield (Bennett and Pratt, 2001).

To further explore the conditions under which attentional effects occur and the potential mechanism(s) underlying these effects, other display variables that are often manipulated in studies of spatial attention are the presence, or absence, of post-stimulus masks, and/or distracter stimuli. In tasks of spatial attention, the mechanisms that have been proposed to account for performance improvements with cueing include *signal enhancement* and *noise reduction* mechanisms. The presence of distracter stimuli in displays, where each distracter is viewed as an independent source of noise, enables the mechanism underlying cueing effects in these displays to be established. Many studies have reported cueing effects in displays containing distracters, and have concluded that a mechanism of distracter exclusion, otherwise known as external noise reduction, is responsible for performance improvements in studies of selective spatial attention (Lu and Doshier, 1998; Shiu and Pashler, 1995). While some researchers insist that attention facilitates performance (accuracy and d' sensitivity) only in conditions where the target stimulus is masked (Shiu and Pashler, 1995; Smith, 2000), or is surrounded by distracters (Smith, 2000) others have found facilitation for unmasked targets and targets presented without distracters (Carrasco, Williams, and Yeshurun, 2002; Henderson, 1996).

Carrasco et al. (2002), in a study on the effect of attention in a visual acuity task, found that cueing improved both accuracy (measured in % correct) and reaction times in the absence of a post-stimulus mask. Carrasco et al. interpreted their results as evidence for a selective attention mechanism that enhances spatial resolution via a signal enhancement mechanism; attention improved visual acuity performance in displays devoid of distracters and post-stimulus masks.

While these different cueing and display conditions enable the effect of attention on spatial cueing tasks to be quantified, the conditions under which attention is manipulated has implications for engaging, and explaining, the mechanism(s) involved in selecting information across the visual field. At present, the diversity of experimental conditions (i.e. different cueing procedures, tasks, use of noise, distracters and/or masks, measuring accuracy and/or response times) used renders the picture still unclear as to the precise mechanisms underlying selective spatial attention. However, the debate within the attention research community is not restricted to the conditions under which attentional effects are observed; another long-lasting debate regards the locus of attention effects: does attention operate early in the processing stream or later?

1.5.2 Early versus late attention effects

The debate regarding the early versus late selection of attention is related less to the timing of processing and more to the processing stage at which attention has an effect. In relation to the early selection of attention, perhaps the best-known theory is Broadbent's (1958) *filter theory*. The basic premise of this theory, as a theory of early selection, is that selection occurs early in the stream of

processing prior to stimulus identification. While all input reaching the sensory system are processed and represented at a physical level of description (e.g. an object's location) only those attended stimuli succeed to the stage of processing at which stimulus identification occurs. The capacity-limited mechanism of selective attention, the selective filter, therefore operates to filter out unattended stimuli, selecting only the attended stimuli for further processing. At the other end of this early/late dichotomy are late selection theories proposed by, amongst others, Deutsch and Deutsch (1968) and Duncan (1980). Late selection theories suggest that selection occur late in the stream of processing following stimulus identification. All sensory inputs are processed and represented at a semantic level of description (e.g. placing the object into a familiar category), under no capacity limitations. Selective processing succeeds this stage, so that only attended stimuli transfer to a further stage of processing at which the stimulus becomes available to awareness and memory, and the ability to make an overt response becomes possible.

The applications of electrophysiological techniques that permit the analyses of attentional effects to be segmented into sensory and decision stages are contributing to the debate regarding the early/late nature of selective attention. Do the mechanisms of attention modulate the selection of input early at a sensory stage in the processing stream or later at a decision stage? An introduction to the use of electrophysiological measures to study selective spatial attention and the subsequent findings of research studying the locus of attention effects follows.

1.5.3 ERP studies of selective spatial attention

The measurement of the brain's electrical activity on the scalp using the electroencephalogram (EEG) has been possible for decades (Berger, 1929; cited by Coles, Gratton, and Fabiani, 1990). The more recent development of time-locking a stimulus event to changes in the electrical activity measured by the EEG has enabled the derivation of the event-related brain potential (ERP), reflecting those parts of the EEG that relate to brain activities that occur in relation to perceptual, motor, or cognitive events. As these brain activities elicit only small potentials in relation to the background EEG, averaging this time-locked event-related activity over many stimulus presentations increases the signal-to-noise ratio of emerging deflections that are associated with early sensory processing, and later decision, motor processes. The ERP is presented on a voltage * time waveform containing a number of positive and negative peaks. Unique peaks, referred to by both their polarity and latency (for example P300 refers to a positive peak with a latency of 300 ms) have been interpreted as physiological markers of different perceptual, motor, or cognitive functions (see Rugg and Coles, 1995 for an introduction to research using ERPs as a tool to investigate a range of cognitive functions). The measurement of ERPs comprising separable deflections reflecting early sensory and later decision, motor-related processes, provides additional information regarding the stages of processing at which attention may modulate. Indeed, the fine-grained temporal resolution obtained by recording ERPs has provided insight into the stage(s) at which experimental manipulations of attention modulate the temporal dynamics of perceptual and cognitive processing (see Eimer, 1998 and Luck, Woodman and Vogel, 2000 for reviews). Accounts of these modulations, and the experimental conditions under which they occur, typically observed in

either the amplitude and/or latency of the ERP components of interest, follows in the next section.

1.5.3.1 Central versus peripheral cueing in ERP studies

In relation to the study of visual attention, the measurement of ERPs has proved fruitful, contributing to long-standing debates surrounding the locus of attentional effects, the range of mechanisms supporting attentional selection, and the temporal resolution of attention. Here, to examine the locus of attention effects, we focus on ERP studies that have examined visual selective attention using central and peripheral cueing methods.

Until very recently, ERP studies using central pre-cues consistently revealed enhancement over occipital brain regions on the first visual evoked responses, the P1, a positive component occurring 80 to 130 ms post-stimulus and the N1, a negative component occurring roughly 140 to 190 ms post-stimulus (Mangun, 1995; Mangun and Hillyard, 1991; Martinez, Anllo-Vento, Sereno, Frank, Buxton, Dubowitz, Wong, Hinrichs, Heinze, and Hillyard, 1999). Typically, (with the exception of Martinez et al. and more recently Doallo, Lorenzo-López, Vizoso, Rodríguez Holguín, Amenedo, Bara, and Cadaveira, 2004), those studies used central cues and employed long SOAs, generally longer than 600 ms, to minimize overlap of cue-related and target-related ERP activity. Martinez et al. (1999) employed SOAs of between 400 and 600 ms, and consistent with other ERP studies employing central cues, Martinez et al. found enlarged P1 and N1 components over occipital sites for attended stimuli, indicating a modulation of activity within the processing stream. It is unclear however from their analysis

whether this effect was observed at each SOA, the shorter SOA of 400 ms and longer at 600 ms, as they collapse their data across SOA (neither do they report any behavioural data analysis for attended versus unattended trials). More recently, Doallo et al. (2004) recorded ERPs to compare the time course between peripheral and central cues at SOAs of 100, 300, 500 and 700 ms and failed to find any P1 enhancement with central cues at any of the SOAs used. Their behavioural data, restricted to analysis of RT, however, suggested a validity effect ($p(\text{valid cue} = .75)$) of the informative central cue such that RTs were faster to a validly cued target than invalid.

The few studies employing peripheral pre-cues, however, provide inconsistent results. Hillyard, Luck, and Mangun (1994), employing SOAs of between 600 and 800 ms, found only N1 enhancement for validly cued trials ($p(\text{valid cue} = .75)$) and failed to find enhancement of the P1. Similarly, Eimer (1994) who used a cue-target SOA of 700 ms, also found N1 enhancement at parietal electrodes for validly cued trials ($p(\text{valid cue} = .75)$), and contrary to those studies using central cues, Eimer actually reported a smaller P1 for valid than for invalid trials. Yet, P1 and N1 enhancement for valid trials have been reported over occipital and temporal sites by Fu, Fan, Chen and Zhuo (2001) who used SOAs from 100 to 300 ms and peripheral pre-cues with a validity of ($p(\text{valid cue}) = .75$).

Using uninformative peripheral pre-cues ($p(\text{valid cue}) = .5$) and long SOAs (566 to 766 ms), Hopfinger and Mangun (1998) also failed to find a reliable P1 enhancement effect; rather at longer SOAs P1 amplitude was reduced using valid cues compared to invalid cues, consistent with Eimer (1994). Yet, in trials with shorter SOAs (34 to 234 ms), Hopfinger and Mangun did find P1 enhancement for validly cued trials over lateral occipital electrode sites (see also Lubbe and

Woestenburg, 1997 who found a P1 enhancement at posterior sites for validly cued trials with SOAs of between 100 and 300 ms). Most recently, comparing SOA with cue informativeness, Doallo et al. (2004) found P1 enhancement with informative ($p(\text{valid cue}) = .75$) and uninformative ($p(\text{valid cue}) = .5$) peripheral pre-cues for their shortest SOA of 100 ms over occipital sites. In contrast, at longer SOAs the reverse P1 amplitude pattern was observed; a smaller P1 for validly than invalidly cued trials was observed with informative peripheral pre-cues at a SOA of 500 ms and at a SOA of 700 ms with uninformative peripheral pre-cues. For the latter cue type this finding extended to SOAs of 300 and 500 ms at parietal electrode sites. No analysis of the N1 was reported in this study.

To summarise, using informative central cues ($p(\text{valid cue}) = .75$ to 1), with the exception of the study by Doallo et al. (2004), enhancements in the amplitude of early ERP components, the P1 and N1, over occipital regions have been consistently observed at long SOAs (from 400 to 800 ms). Using shorter SOAs and central cues, no P1 enhancement for valid cues has been reported (Doallo et al., 2004). For informative peripheral cues, P1 and N1 enhancements have been observed when using SOAs from 100 to 300 ms (Fu et al., 2001), and at 100 ms SOA for uninformative peripheral cues ($p(\text{valid cue}) = .5$). For peripheral cues, at longer SOAs, the effect seems to reverse; P1 and N1 amplitude increases for invalid peripheral cues relative to valid cues at a SOA of 300 ms with uninformative cues, and at 500 ms with informative cues. While such effects may seem reflective of neuronal refractoriness between the cue and the target, Hopfinger and Mangun (1998), argue that neuronal refractory effects would be evident only at short rather than long SOAs.

The effect of cue validity on the P1 component expressed as an amplitude enhancement has been interpreted in terms of a mechanism of sensory gain control, or amplification, that serves to modulate neuronal activity such that attended information elicits larger sensory-evoked responses than unattended stimuli (Hillyard, Vogel, and Luck, 1998). A mechanism of sensory gain control invoked as a mechanism underlying attention-related enhancements of early sensory-evoked brain potentials has also been considered as one of the mechanisms held responsible for performance improvements in a range of visual attention behavioural tasks. However, recent behavioural evidence by Lu, Lesmes, and Doshier (2002) who measured performance for discriminating stimuli in varying levels of external noise suggest that performance improvements in attended locations are enabled by a mechanism that operates to suppress external noise. More work is necessary in ERP research to ascertain the mechanisms behind such enhancements.

In relation to the time course of central and peripheral cueing, Eimer (2000) compared the time course of informative central and peripheral cues at short (200 ms) and long (700 ms) SOAs. In addition to revealing enhanced negativity for validly cued trials at both short and long SOAs, with peripheral cues the effect commenced around 150 ms post-stimulus, but was only later in the central cueing conditions, evident at 250 ms. This latency difference between the two cue types supports behavioural evidence suggesting that peripheral cues invoke an exogenous mechanism characterised by a fast transient response, whereas central cues invoke the slower sustained response of the endogenous system.

Whereas spatial attention has been found to affect the amplitude of early sensory ERP components like the P1 and N1, what is the effect of attention on the

latency of these early sensory-related components? If a function of attention is to prioritise and select task-relevant information from the sensory input, a shortening of the latency of these early sensory-related components may also be beneficial to a cognitive agent in a dynamic environment. Thus far however, P1 and N1 latencies have been found to be largely unaffected by validity. The study of Fu et al. (2001) actually reports longer P1 and N1 latencies for valid than invalid trials. So far, contrary to behavioural evidence, there is little ERP evidence to support a speed-up of early sensory processing for attended as compared to unattended information. Only one report (Di Russo and Spinelli, 2002), using the steady-state visual potential (SSVP) over parietal-occipital areas in a task of sustained, voluntary attention, revealed a reduction of latency by 28 ms between attended and unattended conditions. The latency value for the attended condition was 122 ms, within the time range where the P1 is typically observed. A latency difference has been observed in the later P300 (or P3) component, a positive deflection typically observed from 300 to 900 ms over parietal/central areas, after task-relevant events. The P300 is considered a marker of stimulus categorisation time (McCarthy and Donchin, 1981; for a review of the P300 see Verleger, 1997), and is largely unaffected by speed-accuracy tradeoffs in RT tasks. The results of Fu et al. (2001) revealed a shorter P300 latency for valid than invalid trials, consistent with the effect of decreasing RT with valid pre-cues (Posner, 1980).

The lack of evidence for an effect of cue validity on the latency of the P1/N1 is surprising considering the vast behavioural evidence supporting a shortening of RT to cued targets, than to un-cued targets. In the following section, I consider issues related to measuring the effect of attention on the temporal dynamics of

processing, both in relation to RT measures and ERP latency. To this end, I describe a procedure designed to provide a conjoint measure of response accuracy and response speed, and review research that has used such procedures to determine the locus of response speed improvements and the effect of attention on the speed of information processing.

1.5.5 Measuring the speed of information processing

Inferences about the relative dynamics of perceptual or cognitive processing are traditionally drawn from measurements of simple or choice reaction times. Yet, a RT difference due genuinely to faster processing is indistinguishable from one that is due, instead, to other factors, sensory- or decision-related. For example, it is usually observed that detection RT in Posner's attentional cueing paradigm is faster for validly cued locations than invalid (Posner, 1980). Does this RT pattern indicate a faster processing speed for attended than unattended items? Not necessarily, the discriminability of a target stimulus can also affect processing time; an above threshold stimulus may reach a response threshold earlier than a threshold stimulus at the level of motor processing, yet maintain similar processing speeds at a sensory/decision stage. Additionally, an observer's decision criteria may alter between valid and invalid trials, in turn affecting reaction times. The interpretation of temporal differences in processing using traditional RT measures is further complicated by studies demonstrating enhanced discriminability with attention (Bashinski and Bacharach, 1980; Carrasco, Penpeci-Talgar, and Eckstein, 2000; Posner, 1980; Yeshurun and Carrasco, 1998, 2000). These problems can be made resolute by employing

speed-accuracy trade-off (SAT) procedures. The use of SAT procedures to examine the dynamics of processing has a long history in psychology tracing back over a century (Wickelgren, 1977).

Speed-accuracy trade-off procedures are simply a set of experimental methods that are employed to encourage observers to trade accuracy for speed of performance in order to obtain a measurement of accuracy at a range of response times (Pachella, 1974; Wickelgren, 1977). An example of such a procedure is the response-signal speed-accuracy trade-off method. This procedure simply forces participants to trade response speed for accuracy by presenting them with a response tone, at variable times after target offset, to which they must give an immediate response (within a time window of 300 ms). By varying the lag between the stimulus offset and the tone, a range of processing times could be sampled. The outcome of such experiments is a speed-accuracy trade-off function that conjointly depicts a measure of discriminability with the speed of information processing for a given task. Figure 1.12 depicts a hypothetical SAT function plotting a measure of accuracy, here d' , against the full time-course of an information-processing task.

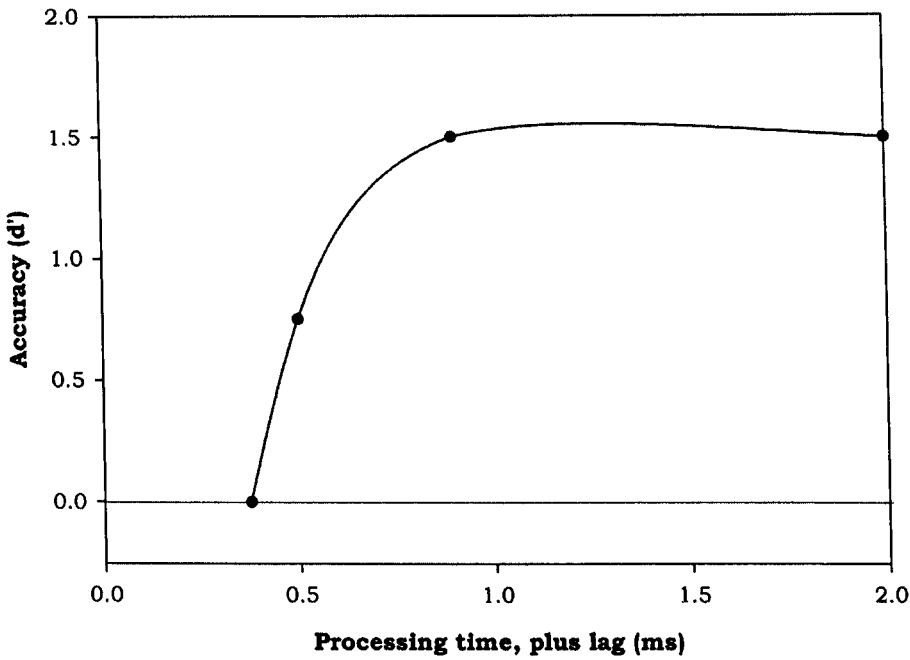


Figure 1.12. A hypothetical speed-accuracy trade-off function. The function depicts conjointly a measure of discriminability (here, d') and processing speed across a full time-course of an information-processing task.

This function can be described with three parameters: an *asymptotic* parameter that reflects saturated performance; a *rate* parameter reflecting the rate of increase in accuracy to asymptote, and an *intercept* marking the point in time where accuracy departs from chance (Wickelgren, 1977).

SAT methods have been used to examine the dynamics of processing in a range of cognitive tasks (e.g. McElree and Carrasco, 1999; Ratcliff, 1978; Reed, 1973). Indeed, McElree and Carrasco (1999; see also Carrasco and McElree, 2001 and Carrasco, McElree, Denisova and Giordano, 2003) used the response-signal SAT procedure to examine the effect of covert attention on the speed of information processing in a visual search paradigm, employing a central neutral cue and a valid peripheral cue ($p(\text{valid cue} = 1)$) in a simple orientation

discrimination task. Figure 1.13 depicts the potential effects of attention on the SAT function.

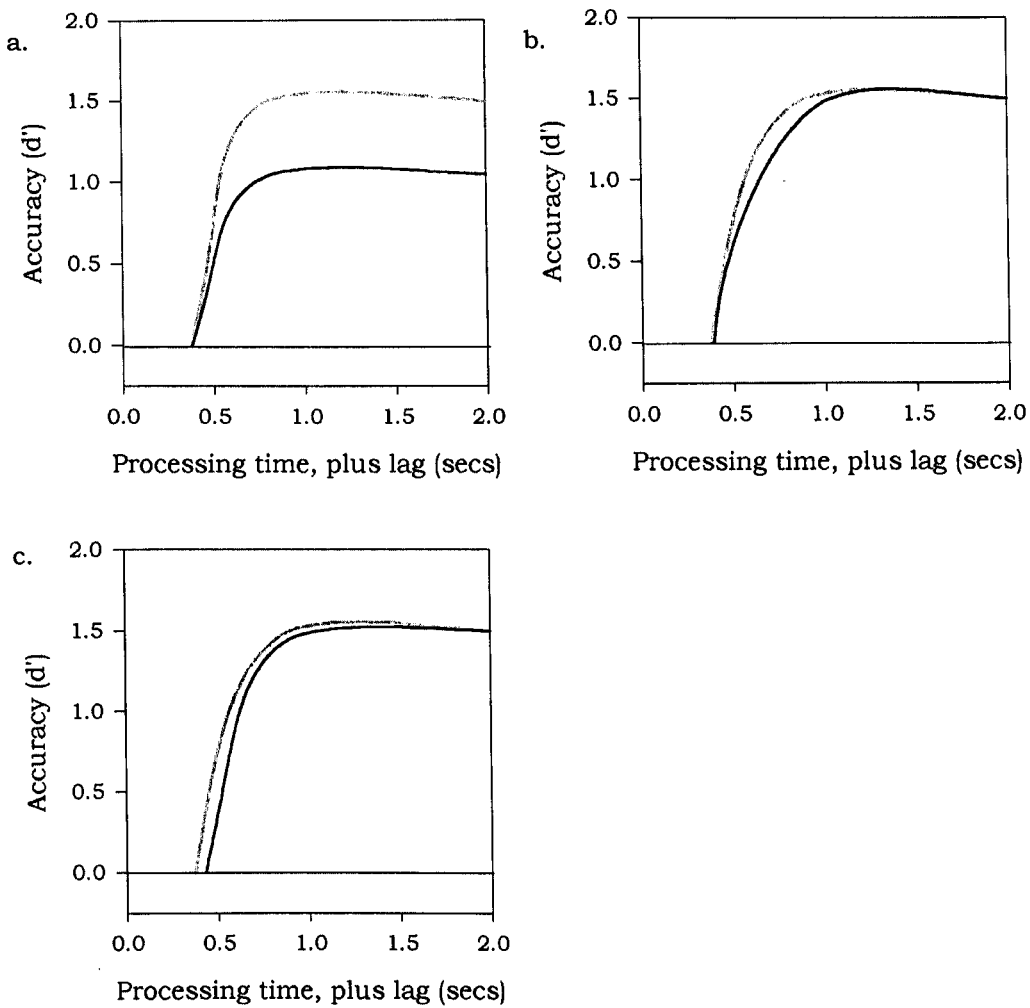


Figure 1.13. Hypothetical SAT functions. Each panel illustrates a potential effect of cueing on the SAT function, plotted in d' units versus processing time (onset of the response tone signal plus response time) in seconds. The plot in panel (a) reveals different performance asymptotes, an expected pattern when attention enhances only target discriminability. Here, attention improves performance accuracy without altering processing time. The plots in (b) and (c) depict two ways in which attention can alter the temporal dynamics of information processing; (b) shows how attention can modify the rate at which information is processed, so that asymptotic performance is reached at a faster rate, whereas (c) depicts how attention can induce an earlier departure from chance performance, reflected in the intercept parameter of the SAT function.

Figure 1.13a depicts the situation where attention modulates the asymptotic parameter, that is improves target discriminability only. An effect on the speed of processing can be expressed either on the rate parameter (Figure 1.13b), or the intercept parameter (Figure 1.13c). A modulation of the rate parameter reflects a faster approach to asymptotic performance, whereas a modulation of the intercept relates to an earlier departure from chance performance.

The results of Carrasco and McElree (2001) are depicted in Figure 1.14 (feature task results shown). The figure shows processing time plotted as a function of accuracy for peripheral valid, and central neutral cues, for different set sizes when the target is presented in isolation or amidst distracter stimuli. Specifically, their results showed a difference in discriminability, measured by asymptotic performance between cued and neutral conditions, a difference of 0.10 d' units in the feature search condition with the smallest set size to 0.55 d' units as the set size increased. Attention, therefore, improved the discriminability of the target signal. In relation to the speed of processing, attention speeded up the rate at which information was processed by 45 ms in the feature search condition, and from 33 (smallest set size) to 106 ms (largest set size) in the conjunction search conditions (not shown here). The results of Carrasco and McElree (2001), employing an SAT method, reveal how attention enhances both the discriminability of a visual input, and the temporal efficiency at which selected visual input is processed.

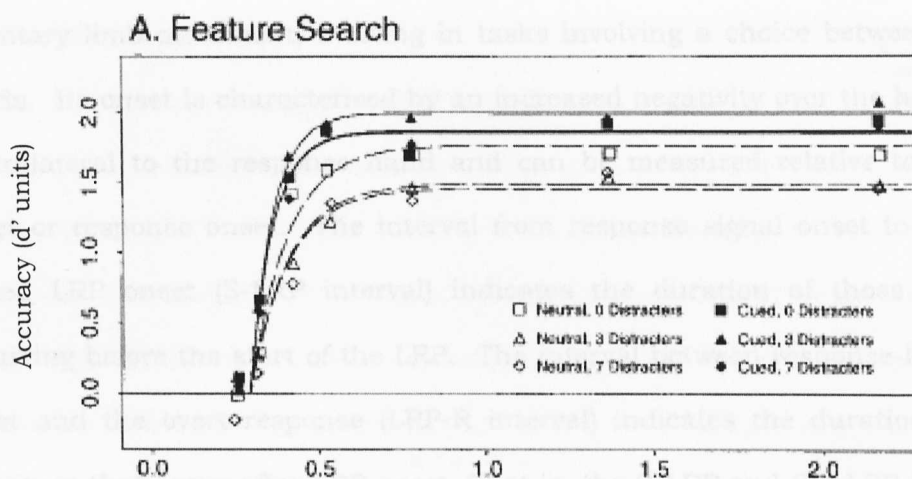


Figure 1.14. The results of Carrasco and McElree (2001). The figure shows SAT functions for cued and neutral trials when the target stimulus is presented alone, and when it is presented amidst distracter stimuli.

A limitation of this research, however, is that the locus of this attentional effect is unknown. In relation to the locus of SAT effects, an outstanding issue concerns the aspect of processing time that is modulated by attention, the pre-motoric or motoric portion. The locus of any attentional effect, pre-motoric or motoric, can be established by examining the lateralised readiness potential (LRP). In the next section, I describe an electrophysiological measure, the lateralised readiness potential (LRP), enabling the partitioning of sensory or pre-motoric processing from decision response-related processing.

1.5.6 SAT and the Lateralised Readiness Potential (LRP)

The LRP reflects an asymmetrical readiness potential, that is, a negative EEG potential that slowly increases and reaches a maximum deflection just before a

voluntary limb movement, evolving in tasks involving a choice between the two hands. Its onset is characterised by an increased negativity over the hemisphere contralateral to the response hand and can be measured relative to stimulus onset or response onset. The interval from response signal onset to stimulus-locked LRP onset (S-LRP interval) indicates the duration of those processes occurring before the start of the LRP. The interval between response-locked LRP onset and the overt response (LRP-R interval) indicates the duration of those processes that occur after LRP onset. That is, the S-LRP and the LRP-R intervals can be used as information processing markers for pre-motoric and motoric processing, respectively. We can therefore analyse LRP onsets obtained from waveforms time-locked either to response signal onset or to the overt response (Osman and Moore, 1993; Leuthold, Sommer, and Ulrich, 1996) to make inferences about the stage of processing that may be modulated by attention.

The lateralised readiness potential has previously been used to establish the locus of speed-accuracy trade-off effects (Rinkenauer, Osman, Ulrich, and Mattes, 2004). Rinkenauer et al. found in a perceptual and cognitive task that the onset latencies of both the S-LRP and the LRP-R were when response speed was stressed. At which stage in processing, sensory or decision-related, does attention speed up information processing? In Chapter Five, I apply the SAT method and orientation discrimination task of Carrasco and McElree (2001) to determine the locus of this attentional modulation. We recorded ERPs while observers completed three spatial attention experiments to assess whether the effects of attention would be observed as a modulation on early sensory-related ERP components, or in later decision-related components.

1.5.7 Summary

To summarise, selective spatial attention has consistently been shown to enhance performance on various visual tasks on both behavioural and electrophysiological measures. These performance improvements are noted as enhanced discriminability of attended visual input, and more recently as acceleration in the processing of selected input. The locus of this effect within the stream of information processing is the object of an ERP study presented in Chapter 5.

1.6 Organisation

The following four chapters present experiments, using a range of methodologies, that address different issues related to the selective use of information for perceptual and recognition tasks.

In Chapter 2, I use the Bubbles technique to determine the information underlying the stable percepts of a bi-stable image, Salvador Dali's *Slave Market with the Disappearing Bust of Voltaire*. On discovering that the information underlying each percept was grounded in different spatial frequency channels processing the image, I used spatial frequency adaptation to selectively induce a perception of the image orthogonal to the adapting frequencies. Taken together, the experiments in Chapter 2 demonstrate the importance of understanding the information underlying a percept and how this information is used for perception.

Since the adaptation method used in Chapter 2 was successful in inducing a selective perception of a complex bi-stable image, I extended the method in

Chapter 3 to attempt an adaptation of local image features (as opposed to adapting globally across the image as in Chapter 2) in Experiment 3.1 again to the Perception of the Dali image, and in Experiment 3.2 to the recognition of gender in hybrid faces. In Experiment 3.1, I adapted observers to the spatial frequency information underlying a percept that was restricted to the spatial regions relevant to that percept, and found again that participants experienced a perception orthogonal to the adapting pattern. Presenting the adapting patterns underlying one percept into the regions of the alternative percept disrupted the pattern of responses previously observed. The results of Experiment 3.1 showed that adapting to the spatial frequency information underlying a percept inhibits its perception only when the adapting pattern is localised to the regions of the image underlying its perception. In Experiment 3.2, I applied the local adaptation method to the perception of gender in hybrid faces that comprised a face of one gender in the diagnostic regions of the face, and a face of another gender in the non-diagnostic regions. In contrast to the results of Experiment 3.1, adapting to the spatial frequencies underlying these faces in either the diagnostic or non-diagnostic regions did not inhibit a gender judgement of the face within that region.

Chapter 4 aims to characterise the mechanism(s) enabling the selective use of visual information, and to localise learning to specific facial regions. Using the method of noise masking in a perceptual learning paradigm, I measured contrast thresholds in varying levels of external noise density, independently across different facial regions, over several days. I compared the information used by human observers to discriminate face stimuli in noise to that of an ideal observer

that could use all of the information available, and found that learning increased efficiency to use information differentially across facial regions for two observers.

The final experimental chapter, Chapter 5, takes a different approach; rather than looking at the information used to resolve a task I examine instead whether or not selective attention to visual information modulates the speed at which this information is processed. In addition to measuring the behavioural response in terms of both accuracy and reaction time using SAT procedures in three experiments, I also measured event-related brain potentials to determine the locus of any effects. Analysis of the behavioural data of three experiments showed that attention modulates the temporal dynamics of visual information processing by enhancing the speed of processing. Analysis of the electrophysiological data did not reveal an effect of attention on the onset latency of the stimulus-locked or response-locked lateral readiness potential. The effects of attention on the amplitude and latency of the ERP waveforms arose as late effects, as opposed to early effects.

A summary of the main aims and findings of each experiment is presented in the final General Discussion of Chapter 6. Here, methodological limitations and theoretical implications are discussed in light of the presented results.

Chapter 2

The Visual Information Underlying the Perception of a Bi-stable Image⁶

The perceptual reversal of bi-stable images has been a source of fascination for psychologists and artists alike, albeit for different reasons. For some artists, the allure of introducing ambiguity is to create in the observer an experience that is, explicitly, purely subjective and qualitative. It is a way of emphasising the constructive nature of perception, the observer's share. For the psychologist, on the other hand, image ambiguity serves as a tool to probe the organisation and dynamics of the visual and cognitive system: the retina receives a single image comprising multiple interpretations, yet the visual and cognitive system is constrained so that only one percept is available at a given time.

⁶ The two experiments reported in this chapter, Experiments 2.1 and 2.2, formed part of a poster presentation at the Annual Conference of the Vision Sciences Society 2001, and is an elaboration of Bonnar L., Gosselin F. & Schyns P.G. (2002). Understanding Dali's *Slave Market with the Disappearing Bust of Voltaire*: A case study in the scale information driving perception. *Perception*, **31**, 683-691.

Classical examples of bi-stable ambiguous figures include those based on figure-ground such as Rubin's vase/faces figure, others perspective-based include the Schroeder staircase and the Necker cube, and object ambiguities like the duck-rabbit figure. The use of such ambiguities is evident in the work of many artists; these 'double-images' were in fact a favourite tool of the Surrealist painter, Salvador Dali (Descharnes, 1972). One such example of Dali's paintings is *Slave Market with the Disappearing Bust of Voltaire* (1940) (see Figure 2.1) in which the heads of two nuns within a busy scene also constitute the eyes of the Bust of Voltaire. On viewing this painting, perception switches from one interpretation to the other.



Figure 2.1. Salvador Dali's *Slave Market with the Disappearing Bust of Voltaire* (1940).

Theoretical explanations for the perception of ambiguous images diverge between theories of inhibition that postulate a satiation in the response of the population of neurons processing the image (Attneave, 1971; Blake, 1989), and attentional theories proposing a perception of bi-stable stimuli based on top-down selection (Leopold and Logothetis, 1999; Meng and Tong, 2004). The former account proposes that activation of sensory neurons processing one image interpretation, and the suppression of neurons responsive to the alternative percept, over time, causes the state of the system to reverse as a result of satiation or fatigue so that the population of neurons responding to the alternative percept become activated, resulting in a switch in the perception of the bi-stable image. The alternative account proposes that perception of bi-stable images is controlled by signals sent from top-down attention-related regions to guide neuronal activity toward a particular interpretation. Alternations between different perceptual interpretations are therefore influenced by factors under the observer's control such as attention and mood, and can be modulated with practice (for reviews on issues related to the perception of bi-stable images, see Blake and Logothetis, 2001 and Leopold and Logothetis, 1999).

Here in Experiment 2.1, we examine the information in an ambiguous figure, Dali's *Slave Market with the Disappearing Bust of Voltaire*, as a case study to determine the information underlying each alternative percept, and in Experiment 2.2 use this information to suppress one percept and induce the other. Our approach which begins with a determination of the information underlying each percept, proposes a switching mechanism that is consistent with theories of inhibition, that is that the population of neurons processing one image

interpretation become satiated, thereby inducing the perception of the alternative image interpretation.

2.1 Revealing the visual information for perception

A general problem in vision that also applies to these bi-stable images is to know which information drives their perception. Ambiguous images, such as Dali's painting, are perfect for investigating the information in a stimulus that underlies its perception; the bottom-up information for each alternative percept is contained in a single image. Until recently, however, no generic technique has been available to simplify a stimulus to the essential information driving its perception. To visualise the information driving the perception of each image interpretation in Dali's painting, we adapt one such technique, *Bubbles* (Gosselin and Schyns, 2001), originally developed to isolate recognition information.

The Bubbles method randomly searches an image generation space to present sparse versions of images as stimuli. Applying a mask, punctured with randomly located Gaussian windows that reveal information, generates each sparse stimulus. By randomly varying the spatial location of each bubble, on a trial-by-trial basis, the entire stimulus is sampled over the course of an experiment. To obtain the information sufficient to resolve a task, the number of bubbles revealing information is adjusted to maintain a pre-specified performance criterion, e.g. 75% correct. Observers respond according to their percept of the sparse stimuli and by retaining the samples of information that lead to each response category the visual information that underlies a percept can be established.

So far, the technique has mainly been applied to recognition tasks involving an objective response. For example, Gosselin and Schyns (2001) used Bubbles to reveal the information that was diagnostic to categorise the identity, gender and expression of the same set of faces. In a between-subjects design, they depicted the diagnostic information for each face categorisation task, demonstrating that the information use of the same face set varied as a function of task demands, across both the 2D image and spatial scales.

Here, we adapt the technique to address a subjective task, the perception of a bi-stable image. To disambiguate the image and to determine the specific visual information that drives each possible percept of Dali's painting we searched the 2D-image as well as the third dimension of spatial scales. We chose to search spatial scales since we know that early vision analyses visual input using multiple spatial filters (see de Valois and de Valois, 1990 for a review, and Chapter 1, Section 1.3). In Experiment 2.1, we found that the nature of the information underlying the alternative percepts in this bi-stable image is grounded in different spatial filters processing the image. In Experiment 2.2, we validated this result using an established psychophysical technique known as frequency-specific adaptation (Blakemore and Campbell, 1969; de Valois and de Valois, 1990). The importance of understanding the information content of a stimulus and how the use of this information underlies perception is highlighted.

2.2 Experiment 2.1

To determine the use of information specific to each stable percept of Dali's painting, we adapted Bubbles (Gosselin and Schyns, 2001). This version of

Bubbles conserves the essential components of the method, as originally conceived, that is, it randomly searches an image space for information underlying perceptual processing, modified slightly for this subjective task (described in the Method section). On presentation of each sparse stimulus, observers respond according to their percept and the samples of information that lead to the 'nuns' response and the 'Voltaire' response are retained. We can establish which visual information selectively drives the percept of each image interpretation.

2.2.1 Method

2.2.1.1 Participants

Participants were ten consenting University of Glasgow students with normal or corrected to normal vision and were paid for their participation.

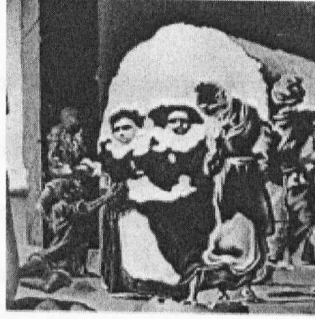
2.2.1.2 Stimuli & Apparatus

The experiment ran on a Macintosh G4 using a programme written with the Psychophysics Toolbox for Matlab (Brainard, 1997; Pelli, 1997) and the Matlab Pyramid Toolbox (Simoncelli, 1997). Using Photoshop® we cropped the ambiguous portion of a grey-scale version of Dali's *Slave Market with Disappearing Bust of Voltaire* so that it comprised the bust of Voltaire and the two nuns (see the top picture of Figure 2.2), subtending 5.72 x 5.72 degrees of visual angle on the screen (for an image of 256 x 256 pixels). To determine the information driving each percept, we used Bubbles.

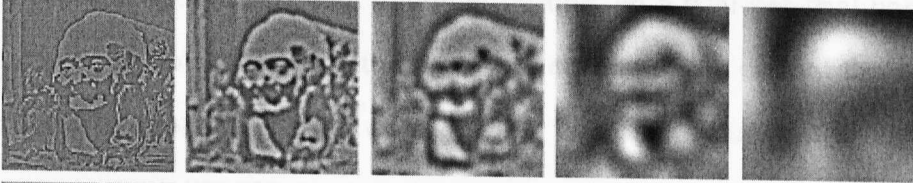
We generated sparse stimuli that randomly sampled an image generation space comprising the 2D-image plane and a range of 6 spatial frequency bandwidths (see Figure 2.2 for a depiction of how the stimuli were generated). These independent bands of spatial frequencies of one octave each had cut-offs at 128, 64, 32, 16, 8 and 4 cycles per image, from fine to coarse scales (in cycles per degree of visual angle, 22.378, 11.189, 5.594, 2.797, 1.399, and .699). The coarsest band was not searched for perceptual information; it served only as a constant background. The image was presented centrally on the screen and the background luminance was 14cd/m². A chin-rest maintained viewing distance at 100 cm.

The image that was represented at each bandwidth was partially revealed by a mid-grey mask that was perforated by a number of randomly located Gaussian windows called bubbles (see Figure 2.2). We normalised to 3 the number of cycles that any bubble could reveal and adjusted the size of the bubble for each frequency band accordingly (the standard deviations of these bubbles were .13, .27, .54, 1.08, and 2.15 cycles per degree of visual angle, from fine to coarse scales). Since the size of the bubbles increases from fine to coarse scales, the number of bubbles at each scale was adjusted to maintain constant, on average, the total area of the image revealed.

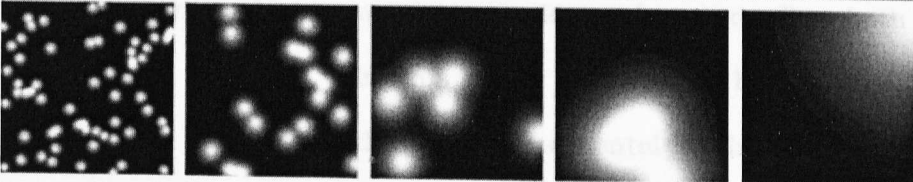
a.



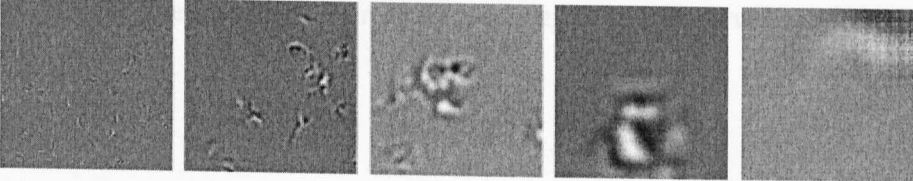
b.



c.



d.



e.

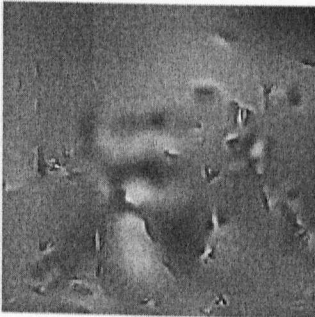


Figure 2.2. The application of Bubbles to the 3D-image generation space. This figure illustrates the method used to apply Bubbles to the 3D image generation space, comprising the 2D ambiguous image and the decomposition into different spatial frequencies. Pictures in (b) represent five different scales of (a); (c) illustrates the bubbles applied to each scale and (d) are the revealed information of (b) by the bubbles of (c). By reconstructing the information in (d) we obtain (e), an example of stimuli presented to observers. Here participants typically report perceiving 'Voltaire'.

2.2.1.3 Procedure

On each trial participants were instructed to indicate by appropriate key presses which image they could perceive, the nuns versus Voltaire. In the event that there was simply insufficient information to perceive either percept, observers were instructed to press a designated 'don't know' key. We introduced the 'don't know' response as a tool to adjust on-line, on a trial-per-trial basis, the total number of bubbles sampling the image generation space so that the number of 'don't know' responses did not exceed 25%. It was emphasised to observers that this response was a last resort and should not be used when both image interpretations were available. In this case, participants had to choose the strongest percept. In each of the 500 experimental trials, a sparse image computed as described appeared on the screen until the observer responded.

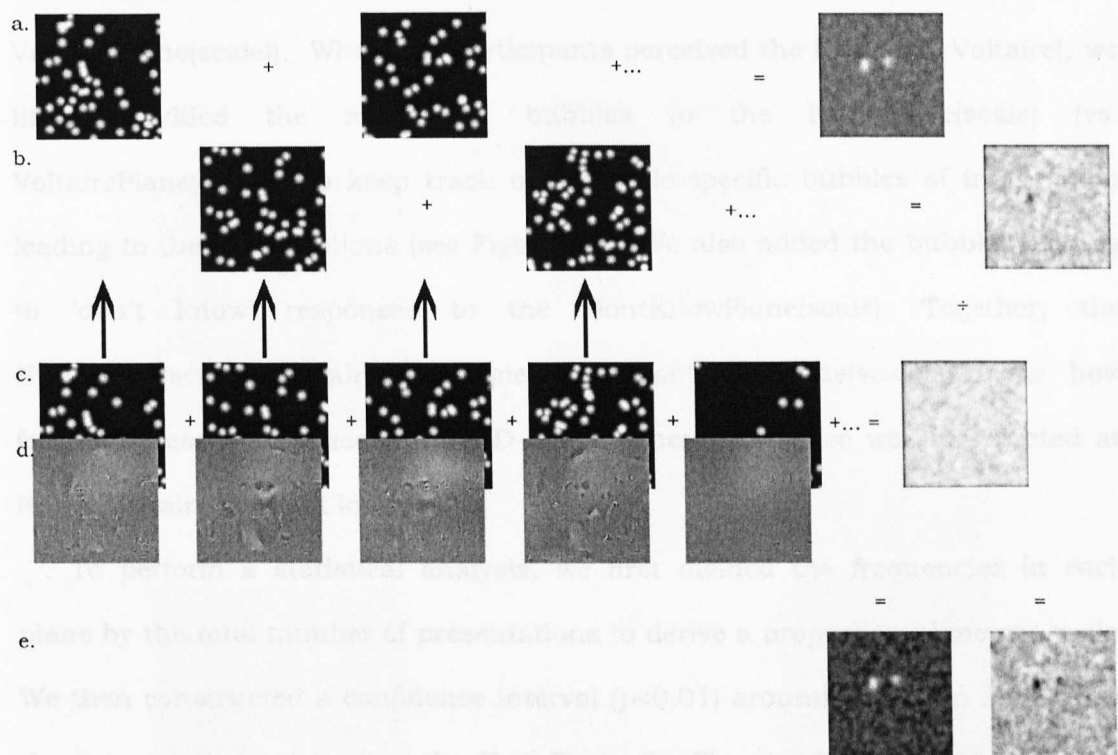


Figure 2.3. Computing the significance of each region in driving each percept. The computations depicted are for the finest scale; identical computations are performed for the four other bandwidths. The pictures in (a) depict the addition of the bubbles that led to a ‘nuns’ response to the NunsPlane, the top right-most picture. The pictures in (b) illustrate the addition of the bubbles that led to a Voltaire response; the right-most picture is the outcome of this addition, VoltairePlane. In (c) all bubbles (those leading to a ‘nuns’, a ‘Voltaire’ and the ‘don’t know’ response) are added to form TotalPlane (the right-most grey-scale picture). Examples of experimental stimuli as revealed by the bubbles of (c) are shown in (d). The pictures in (e), from left to right, are NunsProportionPlane and VoltaireProportionPlane, the division of NunsPlane by TotalPlane and VoltairePlane by TotalPlane. (Note the whiter area in the NunsProportionPlane corresponding to the heads of the two nuns.)

2.2.2 Results

Across trials, we kept track of the locations of the bubbles that led to each image interpretation, the Nuns and Voltaire. To this end, we created a different NunsPlane per scale (for each one of the five scales, henceforth, NunsPlane(scale)), and a different VoltairePlane per scale (henceforth, VoltairePlane(scale)). Whenever participants perceived the Nuns (vs. Voltaire), we literally added the masks of bubbles to the NunsPlane(scale) (vs. VoltairePlane(scale)), to keep track of the scale-specific bubbles of information leading to these perceptions (see Figure 2.3). We also added the bubbles leading to ‘don’t know’ responses to the DontKnowPlane(scale). Together, the NunsPlane(scale), VoltairePlane(scale) and DontKnowPlane(scale) encode how frequently each subspace of the 3D-image generation space was interpreted as Nuns, Voltaire or ‘don’t know’.

To perform a statistical analysis, we first divided the frequencies in each plane by the total number of presentations to derive a proportion plane per scale. We then constructed a confidence interval ($p < 0.01$) around the mean for each of these proportions at each scale, NunsProportionPlane(scale) ($M = .341$, $SD = .092$) and for VoltaireProportionPlane(scale) ($M = .425$, $SD = .078$). We disregarded all of the areas below this confidence interval, as they were considered unimportant in driving each percept. Figure 2.4 depicts the selectively attended information driving each stable percept, that is, those regions above the confidence interval, of the ambiguous portion of Dali’s painting (see Figure 2.1). The top (vs. bottom) pictures depict the information driving the nuns (vs. Voltaire) percept. Note that the nuns occupy the finest and second-to-finest scales with no relevant information at the remaining three coarser scales, and comprise the heads of the

two nuns. In contrast, Voltaire dominates the second, third and fourth scales, with the coarser scales encompassing the entire bust. Clearly, the information driving the two image interpretations differs across the 2D-image plane and across different bandwidths of the spatial spectrum, with an overlap at the second scale.

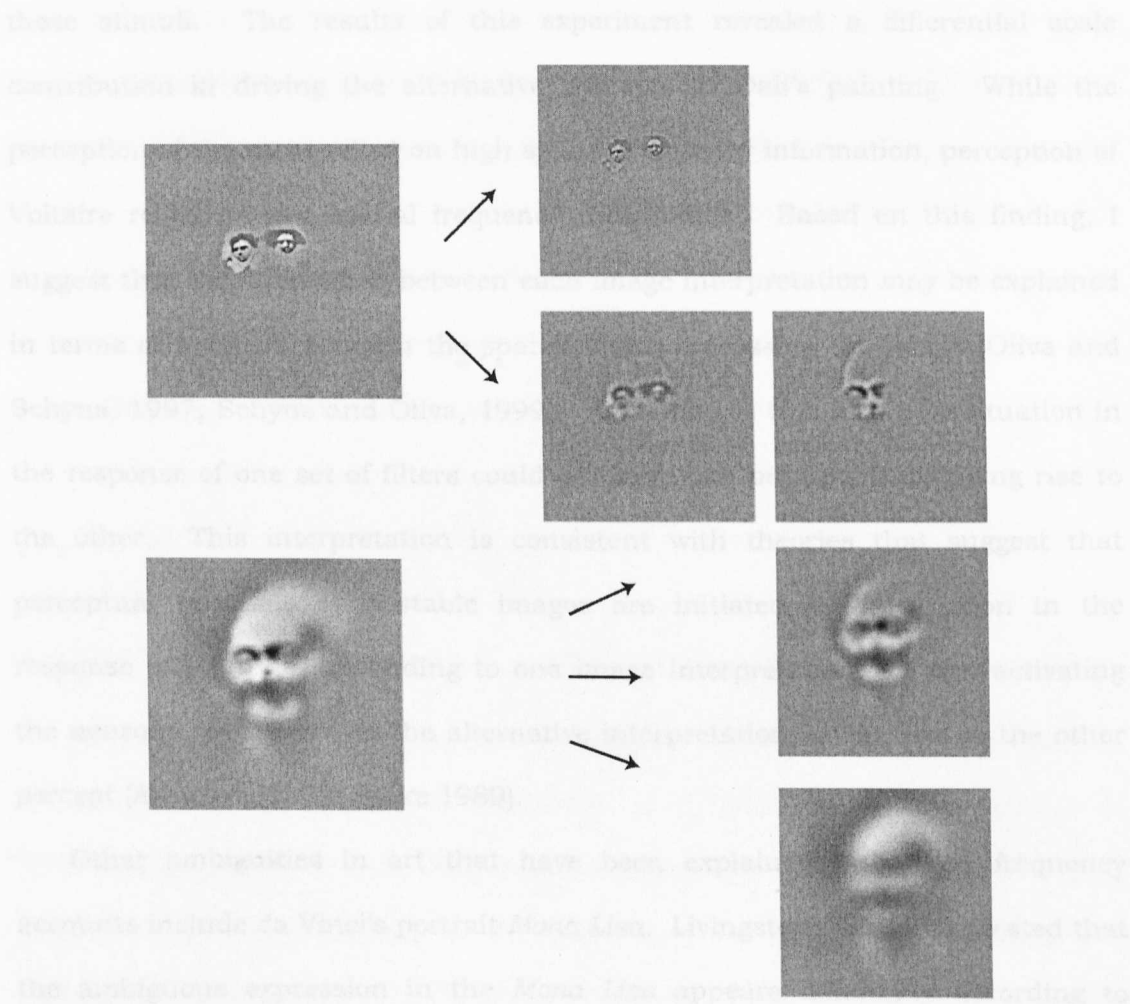


Figure 2.4. The information driving the stable percepts of Dali's *Slave Market with the Disappearing Bust of Voltaire*. The potent information that drives the perception of the 'nuns' and 'Voltaire' is depicted in the left-most column. The top left-most image depicts the potent information driving the 'nuns' percept. The two right-most columns are decompositions of this information into the four scales found to be diagnostic, from finest at the top to the fourth scale at bottom. The second column demonstrates the information relevant to the nuns, and the third column, the information relevant to Voltaire.

2.2.3 Discussion

The aim of Experiment 2.1 was to delineate the information underlying the two stable percepts in Dali's *Slave Market with the Disappearing Bust of Voltaire*. Using the bubbles method, sparse samples of the ambiguous portion of Dali's painting were presented to observers, who responded according to their percept of these stimuli. The results of this experiment revealed a differential scale contribution in driving the alternative percepts in Dali's painting. While the perception of the nuns relied on high spatial frequency information, perception of Voltaire relied on low spatial frequency information. Based on this finding, I suggest that the alternation between each image interpretation may be explained in terms of a switch between the spatial filters processing the image (Oliva and Schyns, 1997; Schyns and Oliva, 1999). According to this idea a habituation in the response of one set of filters could diminish one percept, thus giving rise to the other. This interpretation is consistent with theories that suggest that perceptual reversals of bi-stable images are initiated by a satiation in the response of neurons responding to one image interpretation, in turn activating the neurons responding to the alternative interpretation, giving rise to the other percept (Attneave, 1971; Blake 1989).

Other ambiguities in art that have been explained by spatial frequency accounts include da Vinci's portrait *Mona Lisa*. Livingstone (2000) suggested that the ambiguous expression in the *Mona Lisa* appears differently according to different foveal eccentricities, her smile being more apparent in low spatial frequencies, that is in the peripheral visual field. Hayes and Ross (1995) have also discussed the efficiency of line drawings in representing complex scenes in

terms of the roles of fine and coarse spatial scales, suggesting two distinct systems for processing fine and coarse scale information.

In Experiment 2.1, *Bubbles* isolated the spatial scale information that selectively drives the perception of an ambiguous figure. The method however relies on the participants' categorisations of sparse samples of information potentially inducing strategies atypical of perception (Murray and Gold, 2004a, b; but see Gosselin and Schyns, 2004 for a reply to this criticism).⁷ To strengthen these results, a more direct link between the isolated scale information and classical mechanisms of perception must be established. To this end, Experiment 2.2 will seek to ground the scale-specific perceptions in early vision, using an established psychophysical technique known as frequency-specific adaptation (e.g., Blakemore and Campbell, 1969; de Valois and de Valois, 1990).

2.3 Validating Experiment 2.1 with spatial frequency adaptation

The rationale of frequency-specific adaptation is that an adaptation to pattern X changes the appearance or sensitivity to X, but not the appearance or sensitivity to pattern Y, indicating that the underlying structures simultaneously process independent aspects of the patterns. For example, Blakemore and Campbell (1969) showed that observers exposed to a sinewave pattern oscillating at, e.g. 5 cycles per degree, subsequently exhibited a reduction in their ability to

⁷ Developments regarding the *Bubbles* method, both in its applications and received criticisms, have advanced subsequent to the application of the method presented here. A fuller account of these developments and their implications for the current research is discussed in the General Discussion in Chapter 6.

perceive contrast at this particular frequency, and neighbouring frequencies. That is, adaptation to a spatial frequency selectively impaired sensitivity to this particular frequency, implying that only this channel was affected (see also Pantle and Sekuler, 1968).

Although spatial frequency adaptation has mostly been applied to simple stimuli such as gratings, Webster and Miyahara (1997) used the adaptation method with more complex patterns, scenes containing multiple spatial frequencies and orientations. In a series of experiments they adapted observers to natural images and subsequently measured the effects of this adaptation on the contrast sensitivity function with sine-wave gratings. They found impaired sensitivity to lower spatial frequencies but not for higher.

In Experiment 2.2 I also use multiple spatial frequency adaptation to validate that the information from Experiment 2.1 does drive the perception of the stable percepts in Dali's painting. I reason that if perception of either the nuns or Voltaire results from a switch between distinct spatial channels, spatial frequency adaptation can be used to selectively turn off the spatial filters underlying one percept (e.g., Voltaire) and effectively disambiguating the image (e.g., to induce a stable perception of the nuns). In contrast, if perception results from a switching mechanism solely under the observer's attentional control, then adapting to the spatial frequency information underlying a percept may not lead to its suppression. It is important to stress that our goal is not to adapt the visual system to a specific percept (i.e., Nuns or Voltaire), but rather to habituate the mechanisms of early vision, the spatial frequency channels that mediate the percept. I adapted observers only to the spatial frequencies underlying each

percept and observed the subsequent effects on the perception of the ambiguous image. Importantly, I do not adapt observers to the percept itself.

The results of Experiment 2.1 provide a complete description of the spatial frequencies that must be adapted to selectively affect the percept of the nuns or Voltaire (we describe the adaptation stimuli in the Method section). It remains a challenge to apply frequency-specific adaptation to figurative stimuli because such stimuli do comprise many spatial frequencies at different amplitudes, orientations and phases, most of which must be adapted to obtain the desired effect.

2.4 Experiment 2.2

In the adaptation phase, one group of participants adapted to high contrast dynamic noise created from the high spatial frequencies driving the perception of the nuns. In the other group, participants adapted to high contrast dynamic noise created from the low spatial frequencies driving the perception of Voltaire. In each group, the noise ensured an adaptation to the specific spatial frequencies underlying the percepts, but not to the percepts themselves. In a transfer phase, both participant groups were exposed to a low contrast version of an ambiguous hybrid image composing both the nuns and Voltaire information derived from Experiment 2.1 (see Figure 2.4). If this information is indeed the information underlying these percepts we should then expect the groups to experience orthogonal perceptions on presentation of the ambiguous image. Specifically, the group adapting to low spatial frequencies should perceive the nuns, and the group adapting to the high spatial frequencies should perceive Voltaire. These

orthogonal perceptions subsequent to frequency channel adaptation would establish a possible link between the perceptual information isolated in Experiment 2.1 and fundamental mechanisms of early vision.

2.4.1 Method

2.4.1.1 Participants

Participants were ten fresh consenting students from the University of Glasgow with normal or corrected to normal vision. Half of the participants were assigned to the high spatial frequency adaptation group and half to the low spatial frequency adaptation group. All of the observers were paid for their participation.

2.4.1.2 Stimuli & Apparatus

The experiment ran on a Macintosh G4 using a programme written with the Psychophysics Toolbox for Matlab (Brainard, 1997; Pelli, 1997) and the Pyramid Toolbox (Simoncelli, 1998). For adaptation, we generated 200 (256 x 256 pixels, subtending 13.69 x 13.69 degrees of visual angle on the screen) white noise fields that were filtered to contain the frequency response profile (for the high-frequency noise, between 32 and 64 cycles per image, for the low-frequency noise, it was between 8 and 32 cycles per image) of each potent image derived from Experiment 2.1. Consequently, the noise patterns inherited their high contrast and random phases from the original white noise stimuli, two important properties for frequency-specific adaptation.

The noise patterns were presented dynamically (i.e., in a movie) on a monitor with a screen refresh rate of 75 Hz on a black background, each pattern presented for 13.3 ms. The presentation lasted for 3 minutes and 20 seconds. We constructed a post-adaptation hybrid image combining the low-contrast potent 'nuns' and the low-contrast potent 'Voltaire', a requirement of frequency-adaptation studies (see Figure 2.5). A chin-rest maintained viewing distance at 40 cm.

'nuns' and the potent 'Voltaire' was presented in the centre of the screen for 1 second. Participants were instructed to say aloud which of the two images they could perceive, the 'nuns' or 'Voltaire', and the experimenter recorded these responses.

2.4.2 Results

Of the test subjects, 10 experienced an adaptation effect. The first is on presentation of the 'nuns' image, which was perceived by 10 subjects who adapted to high spatial frequency noise. The second is on presentation of the 'Voltaire' image, which was perceived by 10 subjects who adapted to low spatial frequency noise. To test the effect of adaptation on perception, we compared the results of the perception test under two conditions: (1) no adaptation, and (2) adaptation. The distribution of responses, under a binomial distribution, determines the probability of observing the result by chance. The probability of $p(5)$ of each perceptual event occurring, showed the results to be significant ($p < 0.01$, one-tailed). Adaptation to the spatial

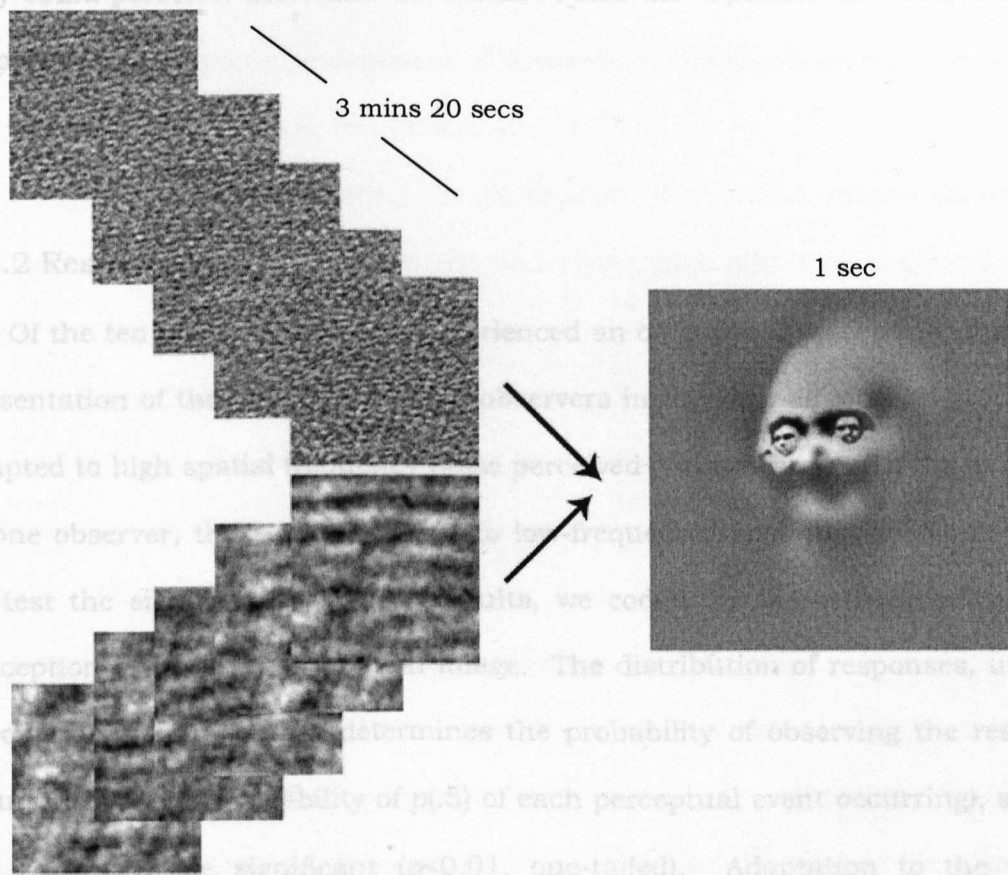


Figure 2.5. Stimuli and design of Experiment 2.2. High-spatial frequency noise with the same amplitude as the 'nuns' is depicted on the top-left and low-spatial frequency noise with the same amplitude as 'Voltaire' is shown on the bottom. Two hundred of these randomly generated noise patterns whose phase was randomly disrupted were repeated 75 times in a movie sequence. The right-most picture is a hybrid comprising the potent nuns and Voltaire information (see Figure 2.2) derived from Experiment 2.1. The contrast of this image was reduced for post-adaptation.

2.4.1.3 Procedure

Before experimentation, to ensure participants could perceive both interpretations of the Dali painting, they viewed the original grey-scale image from Experiment 2.1. Next participants were instructed to adapt to the moving noise patterns by moving their eyes over the pattern (Blakemore and Campbell, 1969). Immediately after adaptation a low-contrast hybrid image comprising the potent 'nuns' and the potent 'Voltaire' was presented in the centre of the screen for 1 second. Participants were instructed to say aloud which of the two images they could perceive, the 'nuns' or 'Voltaire', and the experimenter recorded these responses.

2.4.2 Results and Discussion

Of the ten participants, nine experienced an orthogonal perception; that is on presentation of the hybrid image, all observers in the Nun-adaptation group who adapted to high spatial frequency noise perceived 'Voltaire' and with the exclusion of one observer, those who adapted to low-frequency noise perceived the 'nuns'. To test the significance of these results, we coded for the orthogonality of the perception of the post-adaptation image. The distribution of responses, under a binomial distribution that determines the probability of observing the result by chance (there is a probability of $p(.5)$ of each perceptual event occurring), showed the results to be significant ($p < 0.01$, one-tailed). Adaptation to the spatial frequency information underlying a percept reliably caused the suppression of the percept, inducing the perception of the alternative percept.

These results demonstrate a possible grounding in the mechanisms of early vision the information underlying the selective perception of the nuns and Voltaire isolated in Experiment 2.1 with *Bubbles*. Together, the results of Experiments 2.1 and 2.2 have implications for the study of recognition and perception. An important issue pertains to the information content of a stimulus and its role in driving these processes. Here, it was shown that the reversal between the two percepts in an ambiguous image could be understood in terms of the different scale information underlying these percepts. It is believed that knowing which information people attend to, to resolve visual tasks, is an important but neglected component of perception. As demonstrated the Bubbles technique can reveal what this information is.

Future studies investigating the perception of bi-stable images should take into account the information content underlying each stable percept and its role in the perceptual process. As there are no indications in the literature on bi-stable images that would suggest a systematic bias towards the preponderance of one percept over the other, I did not empirically establish the preponderance of each percept in Dali's ambiguous *Slave Market with the Disappearing Bust of Voltaire*. Nevertheless, the probability of each perception occurring may not be $p(.5)$ as there may be a dominating image interpretation. Consequently, a problem arises for theories of inhibition that propose a satiation in the responses of neurons processing a percept, presumably an extended time spent experiencing one percept as opposed to another should initiate the alternative percept, yet with different response distributions that may be unexplained by such theories. In a related vein, we did not specifically measure the perceptual reversal between the different image interpretations; rather we examined only the

information underlying each stable image interpretation and subsequently used this information to suppress one of the available percepts. The evidence from Experiment 2.2 suggests that this information did suppress the related percept, leading to the speculation that a switch between the different spatial filters may be the mechanism by which perceptual reversal of the interpretations for this particular bi-stable image may be achieved. To investigate this hypothesis, future research should study the process of perceptual reversal explicitly, perhaps focusing on the temporal dynamics of switching between percepts when the observer free-views the original image.

2.5 Concluding remarks

In Experiment 2.1, using Bubbles, we isolated the information that drives the perception of the two stable images of a single bi-stable image. Using an adapted version of Gosselin and Schyns (2001) Bubbles technique to resolve a subjective task, we disambiguated the information for perception and suggested that the reversal between the two image percepts is the result of a switch in the spatial filters processing the image. In Experiment 2.2, participants adapted to low- or high-spatial frequency noise to block the response of that frequency channel and henceforth, upon presentation of a hybrid image, experienced an orthogonal perception. Theoretically, we highlight the importance of understanding the information contained in a stimulus, and how the nature of this information influences perceptual processes.

In the chapter to follow, I consider how the nature of *local* stimulus information influences perceptual processes. I apply the frequency-specific

adaptation method again to the ambiguous portion of Dali's painting, but restrict it to local regions of the image in order to assess the role of local stimulus information. The method is then extended to the perception of gender in hybrid faces to examine the effect of suppressing local stimulus information on participants' gender categorisations.

Chapter 3

Perceiving and Recognising Local Image Features⁸

A generic problem in vision is to know which information drives the perception of a stimulus. This problem was addressed in the case study described in Chapter 2 where we depicted the information driving the two stable percepts of the ambiguous portion of Salvador Dali's painting *Slave Market with the Disappearing Bust of Voltaire* (1940). I applied the psychophysically robust technique of spatial-frequency adaptation to the perception of a complex stimulus comprising multiple spatial frequencies at many orientations. This method also offers promising avenues to diverse problems of information use in perception and recognition, in regards to the role of local features in perceptual processing. Such insights would require, however, an adaptation effect at a local level.

In relation to the problem of information use in perception and recognition, methods such as bubbles are proving to be informative in revealing the subset of a stimulus used by an observer to perform a particular task. The potential role of

⁸ Experiment 3.1 formed part of the poster presentation at the Annual Conference of the Vision Sciences Society 2001, Bonnar, L., Gosselin, F., & Schyns, P.G. (2002). Revealing and suppressing the visual information for recognition. *Journal of Vision*, **2**, 339a.

the adaptation method may be in elucidating the effects of suppressing this information on perceptual processing. In considering adaptation as a tool to suppress the mechanisms responsible for processing aspects of a stimulus we conducted two experiments using a local adaptation method to test the relevancy of local stimulus features in a perception and recognition task. To determine the effectiveness of this method, Experiment 3.1 extends the findings of the global adaptation method of Experiment 2.2 by locally suppressing the mechanisms responsible for processing the alternative percepts of Dali's painting. Experiment 3.2 aimed at further developing the method in applying it to the task of gender recognition of hybrid faces.

3.1 Experiment 3.1

In the current experiment, I attempt to adapt locally the information mediating the stable percepts of Dali's painting, that is, the relevant spatial region as well as the relevant spatial frequencies. To test the interaction between the local region and the spatial frequencies in driving each percept we ran two conditions. In the first condition the relationship between the potent region and the potent spatial frequencies is preserved, consistent with the procedure used in Experiment 2.2, with the exception that here I restrict the spatial frequency adaptation patterns to the spatial region of each stable percept. This first condition will be referred to as the Congruent-adaptation condition. Consistent with adaptation procedures, I observed the effects of the adaptation on the perception of the low-contrast hybrid image. If local adaptation affects the

perception of the post-adaptation stimulus, I would expect a perception orthogonal to the adapting frequency, similar to those observed following global adaptation. Qualitatively, orthogonal responses mean that on presentation of the post-adaptation stimulus, observers adapting to high-spatial frequencies should perceive Voltaire, while those adapting to low-spatial frequencies should perceive the nuns.

In the second condition, I put out of phase the spatial region and spatial frequencies; the high frequency noise in the 'nuns' region was replaced with the low spatial-frequency information underlying the 'Voltaire' percept (and vice versa). The second condition will be referred to as the Incongruent-adaptation condition. If the perception of these image interpretations relies on the interaction between the spatial frequency information and its potent region, a disruption of this information should result in a different distribution of responses than those expected for the Congruent-adaptation condition under the binomial distribution. The distribution of responses following these adaptation conditions can be predicted following consideration of the spatial extent and overlap between the two percepts. Figure 3.1 depicts the spatial extent and overlap between the nuns and Voltaire percepts; the percept of the nuns constitutes a part of the Voltaire percept, the eyes of Voltaire and the spatial extent of Voltaire is greater of that of the nuns. Qualitatively, the impact of adapting to low spatial frequencies restricted to the nuns region should not inhibit a Voltaire response, as there may be sufficient perceptual information to perceive Voltaire beyond the adapting region of the nuns. In contrast, adapting to high spatial frequencies restricted to the Voltaire region should still inhibit a

Nuns response due to the spatial overlap between the nuns region and the Voltaire region.

3.1.1 Method

3.1.1.1 *Participants*

Participants were forty students from the University of Glasgow with normal, or corrected to normal vision, who had not taken part in the previous experiment. Twenty participants were randomly assigned to one of the two conditions (Congruent-adaptation and Incongruent-adaptation) and within each condition half of the participants were assigned to the 'Voltaire' adaptation group ($n = 10$) and half to the 'Nuns' adaptation group ($n = 10$).

3.1.1.2 *Stimuli & Apparatus*

The experiment ran on a Macintosh G4 using a programme written with the Psychophysics Toolbox for Matlab (Brainard, 1997; Pelli, 1997) and the Pyramid Toolbox (Simoncelli, 1998). Using a pseudo-random number generator in MatLab, we created, for adaptation, 200 white noise fields (256 x 256 pixels) that were filtered to contain the frequency response profile of each potent image, derived from the previous experiment (for the high-frequency noise, between 32 and 64 cycles per image, for the low-frequency noise, it was between 8 and 32 cycles per image). In the Congruent-adaptation condition, to restrict the noise only to the important regions underlying each percept, we multiplied the noise fields by the potent masks for each percept, derived from Experiment 2.1. For the Incongruent-adaptation condition, we inserted the high-frequency noise of the

nuns into the potent region for the Voltaire percept and the low-frequency information of the Voltaire percept into the nuns region (see Figure 3.1).

The noise patterns were presented dynamically (i.e., in a movie) on a monitor with a screen refresh rate of 75 Hz (each pattern was presented for 13.3ms). The adaptation stimuli were presented on a grey background for 3 minutes 20 seconds (the sequence of 200 noise patterns was repeated 75 times). We used the post-adaptation hybrid image from Experiment 2.2 combining the low-contrast potent 'nuns' and the low-contrast potent 'Voltaire'. A chin-rest maintained viewing distance at 40 cm.

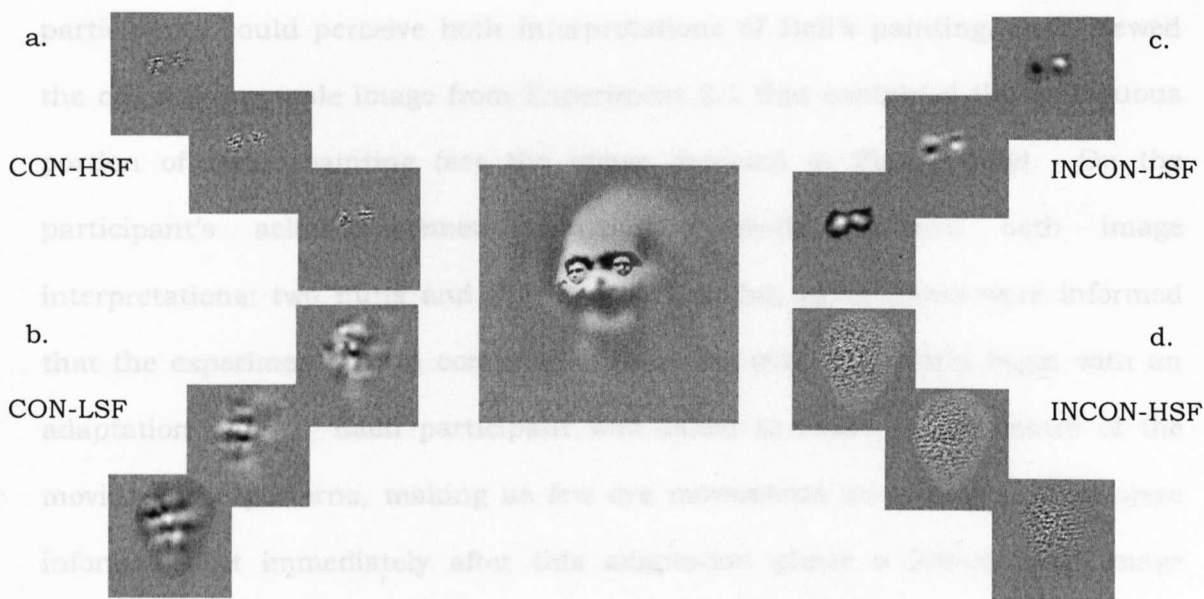


Figure 3.1. The stimuli and design of Experiment 3.1. The left side of the figure depicts the Congruent-adaptation condition. In (a), the upper left quadrant, the high spatial frequency noise with the identical amplitude as the ‘nuns’ is restricted to the potent region of the ‘nuns’ (the CON-HSF condition), and in (b), the lower left quadrant, the low spatial frequency noise of ‘Voltaire’ confined to the ‘Voltaire’ region (the CON-LSF condition). The right side of the figure shows examples of the stimuli used in the Incongruent-adaptation condition. Depicted in the upper right quadrant (c) is the low spatial frequency noise restricted to the nuns region (the INCON-LSF condition) and in (d) the lower right quadrant, is high spatial frequency noise confined to the Voltaire region (the INCON-HSF condition). The image shown in the centre is the post-adaptation hybrid image containing the potent nuns and potent Voltaire information (depicted separately in the left-hand side of Figure 2.4) derived from Experiment 2.1 and depicted in Figure 2.5 (the actual experimental image was of lower contrast than that presented here).

3.1.2 Results

3.1.2.1 Congruent-adaptation

3.1.1.3 Procedure and Design

In a between-subjects design, ten participants were randomly assigned to each of one of the four experimental conditions: the Congruent-adaptation conditions CON-HSF and CON-LSF, where the spatial region and spatial frequency information underlying a percept are preserved, and the Incongruent-adaptation conditions, INCON-HSF and INCON-LSF, where the spatial region and spatial frequency information are disrupted. Prior to the task, to ensure all

participants could perceive both interpretations of Dali's painting, each viewed the original greyscale image from Experiment 2.1 that contained the ambiguous portion of Dali's painting (see the image depicted in Figure 2.2a). On the participant's acknowledgement that they could perceive both image interpretations: two nuns and the Bust of Voltaire, participants were informed that the experiment would comprise of only one trial and would begin with an adaptation phase. Each participant was asked to fixate in the centre of the moving noise patterns, making as few eye movements as possible. They were informed that immediately after this adaptation phase a low-contrast image would be presented for 1 second. Participants were instructed to say aloud which of the two images they could perceive, the 'nuns' or 'Voltaire'. The post-adaptation image that was presented, a low contrast version of the central image depicted in Figure 3.1, was a hybrid image comprising the potent nuns and Voltaire information, derived from Experiment 2.2.

3.1.2 Results

3.1.2.1 *Congruent-adaptation*

Of the ten observers who adapted to the high-spatial frequency information restricted to the potent region of the 'nuns' region (in Figure 3.2 this condition is referred to as CON-HSF), an orthogonal result, that is the perception of Voltaire following adaptation, was evident in 9/10 observers. Of the 10 observers who adapted to the low-spatial frequency information underlying the 'Voltaire' percept, located in the potent 'Voltaire' region (CON-LSF), 9 observers reported perceiving the nuns on presentation of the ambiguous image, a perception orthogonal to the

information to which they adapted. These distributions of responses were both significant under the binomial distribution ($p=0.01$, one-tailed). This result is consistent with the global adaptation effect found in Experiment 2.2, but more importantly, it demonstrates the transfer of a local adaptation effect to the local information within a complex image. Still, it remains necessary to further test the robustness of the interaction between the spatial frequency information and the location of this information in driving perception. The second condition tested the local nature of this information by making incongruent the spatial region and the spatial frequency information underlying each percept.

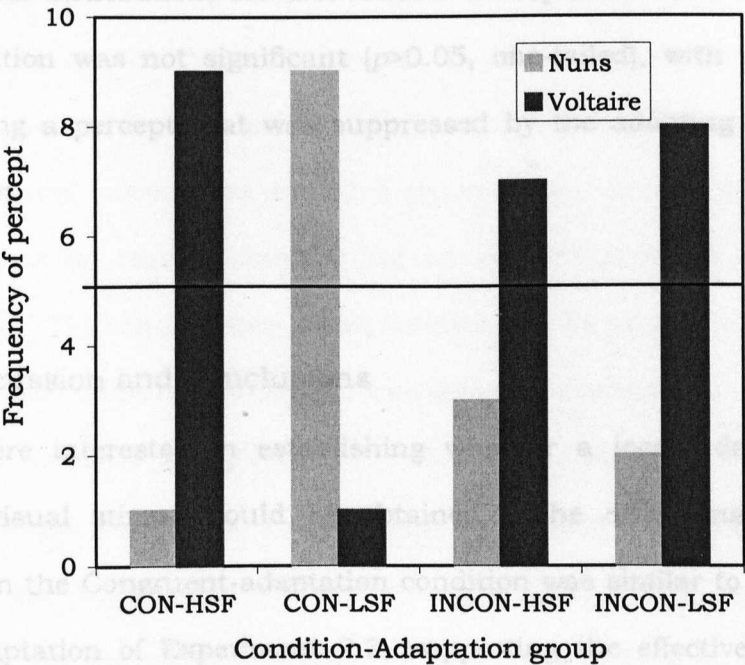


Figure 3.2. The results of Experiment 3.1. The bar chart depicts the frequency of each percept for each of the four conditions. The conditions CON and INCON refer respectively to the preservation of congruency between the spatial region and the spatial frequencies of the adaptation group, and the disruption of this relationship. The adaptation terms HSF and LSF refer to high spatial frequencies and low-spatial frequencies, respectively. Each adaptation pattern encompasses an entire spatial frequency description of either the nuns or of Voltaire.

3.1.2.2 *Incongruent-adaptation*

On adapting to the high spatial frequency noise in the Voltaire region (INCON-HSF), the presentation of the hybrid image resulted in 7 observers perceiving Voltaire and 3 perceiving the nuns. Of the 10 observers who adapted to the low-spatial frequency noise inserted in the nuns region (INCON-LSF), only 2 observers experienced a perception orthogonal to the adapting frequency, that is, 2 perceived the nuns and 8 Voltaire. The distribution of responses for each condition under the binomial distribution proved significant in the INCON-LSF case ($p < 0.05$, one-tailed), with 8/10 participants experiencing a percept that was unsuppressed by the spatial frequency information to which they adapted. Under the binomial distribution, the distribution of responses observed in the INCON-HSF condition was not significant ($p > 0.05$, one-tailed), with 7/10 participants experiencing a percept that was suppressed by the adapting spatial frequency information.

3.1.3 Discussion and Conclusions

We were interested in establishing whether a local adaptation effect on complex visual stimuli could be obtained. The orthogonality of responses observed in the Congruent-adaptation condition was similar to the results of the global adaptation of Experiment 2.2, supporting the effectiveness of the local adaptation. The orthogonal result observed in the Congruent-adaptation condition, where the relationship between the potent region and spatial frequencies was preserved, clearly diminishes when this information is put out of phase in the Incongruent-adaptation condition. The results of the Incongruent-

adaptation condition demonstrate the dependence between the spatial extent and the spatial frequency information driving each percept. That is, adapting to the spatial frequency information of a percept that was not restricted to its spatial region did not necessarily suppress its perception as in the INCON-LSF condition. The local/global nature of the stimulus may explain this result; the information comprising the 'nuns' also constitutes the eyes of Voltaire, thereby inserting high spatial frequencies into the 'Voltaire' region naturally includes inserting it into the 'nuns' region, therefore inhibiting a 'nuns' perception (only 3/10 perceive Voltaire). In inserting the low-spatial frequencies into the 'nuns' region, sufficient 'Voltaire' information is preserved outside the adapting region to perceive 'Voltaire' (8/10). The frequency of responses, as noted in Figure 3.2, suggests a predominance of the 'Voltaire' percept in the incongruent conditions. The observed 'perceptual suppression' demonstrates the importance of the interaction between spatial region and spatial frequencies in driving these two percepts. Altogether, these results demonstrate an adaptation effect on local stimulus information. The results suggest the method may be used as a tool to investigate the role of local stimulus features in mediating perception and recognition.

Further interesting applications of this local adaptation method include the extent to which attentional modulations of contrast determine perception and also, how the information content of a stimulus relates to conscious visual perception. In the section to follow, I apply the method to the recognition of gender in hybrid faces.

3.2 Locally adapting information for recognition

The diagnostic recognition framework (Schyns, 1998) posits a flexible usage of visual information that is constrained by the availability of information, the constraints of the visual system and the demands of the task. In three categorisation tasks (identity, gender and expression) of the same face stimuli, Schyns, Bonnar and Gosselin (2002) demonstrated a differential use of information to resolve these tasks, both within each task (across different spatial scales) and between each task (across the 2D image plane and across spatial scales). In other words, different information in the same stimulus set was diagnostic to resolve the three tasks, demonstrating a flexible use of information as a result of task demands. For example, in judging the gender of a face, the left side of the face was found to be more diagnostic than the right side, whereas, in judging whether a face is neutral or happy, the mouth area alone proved sufficient (Gosselin and Schyns, 2001; Schyns et al., 2002, 2003).

The local nature of these diagnostic regions in resolving a particular categorisation task warrants the validity of applying the local adaptation method used in Experiment 3.1 to observe the effect of suppressing this information on the recognition of the post-adaptation stimulus. We were interested in suppressing the diagnostic information underlying a specific face categorisation to determine whether the classification could still be made when the necessary diagnostic information to make the categorisation was perceptually suppressed. To this end, we created novel hybrid stimuli that contain the information from one face in the diagnostic regions for a particular task and the information from another face in the non-diagnostic regions. For our purposes, the task of gender recognition seemed appropriate due to the asymmetry in selecting facial

information on the left side of the face to resolve gender identification (Burt and Perrett, 1997; Schyns et al., 2002). This asymmetrical processing of visual information is due to the role of the right-hemisphere (attending to the left-visual field) in processing facial gender (Burt and Perrett, 1997; Butler et al., 2005).

3.3 Experiment 3.2

Hybrid stimuli have been used previously to study the flexibility of scale information in the categorisation of faces (Schyns and Oliva, 1999) and scenes (Oliva and Schyns, 1997). However, in these studies observers were sensitised to either low, or to high, spatial frequency faces (or scenes), then categorised hybrid images that combined, for example, an angry man in low-spatial frequencies and a happy woman in high-spatial frequencies. In these experiments, observers were unaware that the stimuli were hybrid stimuli and significantly judged the stimulus consistent with their spatial frequency sensitisation. That is, those observers who were sensitised to high spatial frequencies categorised the hybrid stimulus according to the face that was represented in high spatial frequencies.

In the experiment to be described, the aim and method are different. Here, we use adaptation to suppress specific perceptual information, unlike Oliva and Schyns (1997) and Schyns and Oliva (1999) who sensitised observers to specific perceptual information. In the adaptation phase, we aim to locally suppress the information underlying the gender recognition of one of the faces comprising the hybrid stimulus, and thus induce a gender judgement of the stimulus that reflects the face in the non-adapting region. The adapting patterns are restricted

either to the diagnostic regions of the face or the non-diagnostic regions. An observer, presented with adapting patterns in the diagnostic regions, should experience a suppression of the facial information located in those regions, that is, the eyes, and mouth regions, with a particular focus towards the left side of the face. It is predicted that a suppression of this information, on presentation of a hybrid face, comprising the facial information of a male face and female face, should report the gender of the face in the non-diagnostic region, where information is not suppressed. Conversely, an observer, presented with adapting patterns in the non-diagnostic regions, should experience a suppression of the facial information located in those regions, that is, the right eye and chin regions, forehead and external shape of the face. Here, it is predicted that a suppression of this information, on presentation of a hybrid face, should report a gender judgement that reflects the face in the diagnostic regions. Again, note that observers are adapted only to the spatial frequency information underlying the stimulus, not to the stimulus itself.

3.3.1 Method

3.3.1.1 *Participants*

Participants were nineteen consenting University of Glasgow students with normal or corrected to normal vision, paid for their assistance. Each participant was allocated into one of the two local adaptation groups, five to the diagnostic adaptation region and thirteen to the non-diagnostic adaptation region⁹. These

⁹ The difference in sample sizes between these two conditions is the result of pilot experiments that demonstrated an effect of adaptation on the gender reported of the post-adaptation hybrid image only in the diagnostic adaptation condition.

two groups will be referred to as the Diagnostic-adaptation group and Non-diagnostic adaptation group, respectively.

3.3.1.2 Stimuli & Apparatus

The experiment ran on a Macintosh G4 using a programme written with the Psychophysics Toolbox for Matlab (Brainard, 1997; Pelli, 1997) and the Pyramid Toolbox (Simoncelli, 1998).

3.3.1.2.1 Diagnostic Hybrids

To generate the diagnostic hybrid faces, which were presented post-adaptation, we used the face set and diagnostic face masks for the gender task of Schyns et al. (2002; shown in Figure 3.3a and b). At each spatial scale we inserted the information from one face into the diagnostic region and information from another face of a different gender into the non-diagnostic region. Figure 3.3 depicts the generation of the diagnostic hybrids. The hybrid stimulus set consisted of eight diagnostic hybrid faces (gender counterbalanced across the diagnostic and non-diagnostic regions so that four faces contained males in the diagnostic region and females in the non-diagnostic regions, and vice-versa) that were generated from a set of twelve faces (seven male, five female).

As is necessitated by adaptation procedures the contrast of the hybrid faces was reduced. The reduction in contrast was not uniform across the entire hybrid face; rather, the contrast was differentially adjusted so that the diagnostic face region was of higher contrast than the non-diagnostic face region. The motivation for this differential contrast adjustment was simply to reduce the bias to the face represented in the non-diagnostic region. This region simply contains more

information (see Schyns et al., 2002), reducing the contrast relative to the diagnostic region therefore counter-balances this bias by reducing the salience of the non-diagnostic region.

3.3.1.2.2 Adaptation Stimuli

To generate the adaptation stimuli, we created 1000 white noise patterns that were coloured with the average spatial frequency spectrum of our stimulus set (computed by taking the fourier transform of each hybrid face in the set of stimuli and computing the mean for the set). For the Diagnostic-adaptation condition, using the diagnostic masks for gender (as derived in Schyns et al., 2002), we inserted into the diagnostic areas at each of the five scales a new coloured noise pattern on every frame (on a monitor with a screen refresh rate of 75 Hz, each frame lasted for 13.3 ms). Figure 3.4 shows the stimuli and design of the adaptation experiment. Prior to each trial we randomly selected from the set of noise patterns a subset for the adaptation period. For the initial adaptation period, we selected 200 of these patterns that we repeated 135 times (for a total presentation time of 360 seconds). Consistent with adaptation procedures, prior to each subsequent trial, an additional 50 adaptation patterns were randomly selected and repeated 45 times (lasting 30 seconds), to maintain the adaptation effect throughout the experiment (see De Valois and De Valois, 1990 for a review of studies and procedures used to adapt spatial frequency channels). A mid-grey fixation cross was placed in the centre of the adapting pattern to maintain fixation.

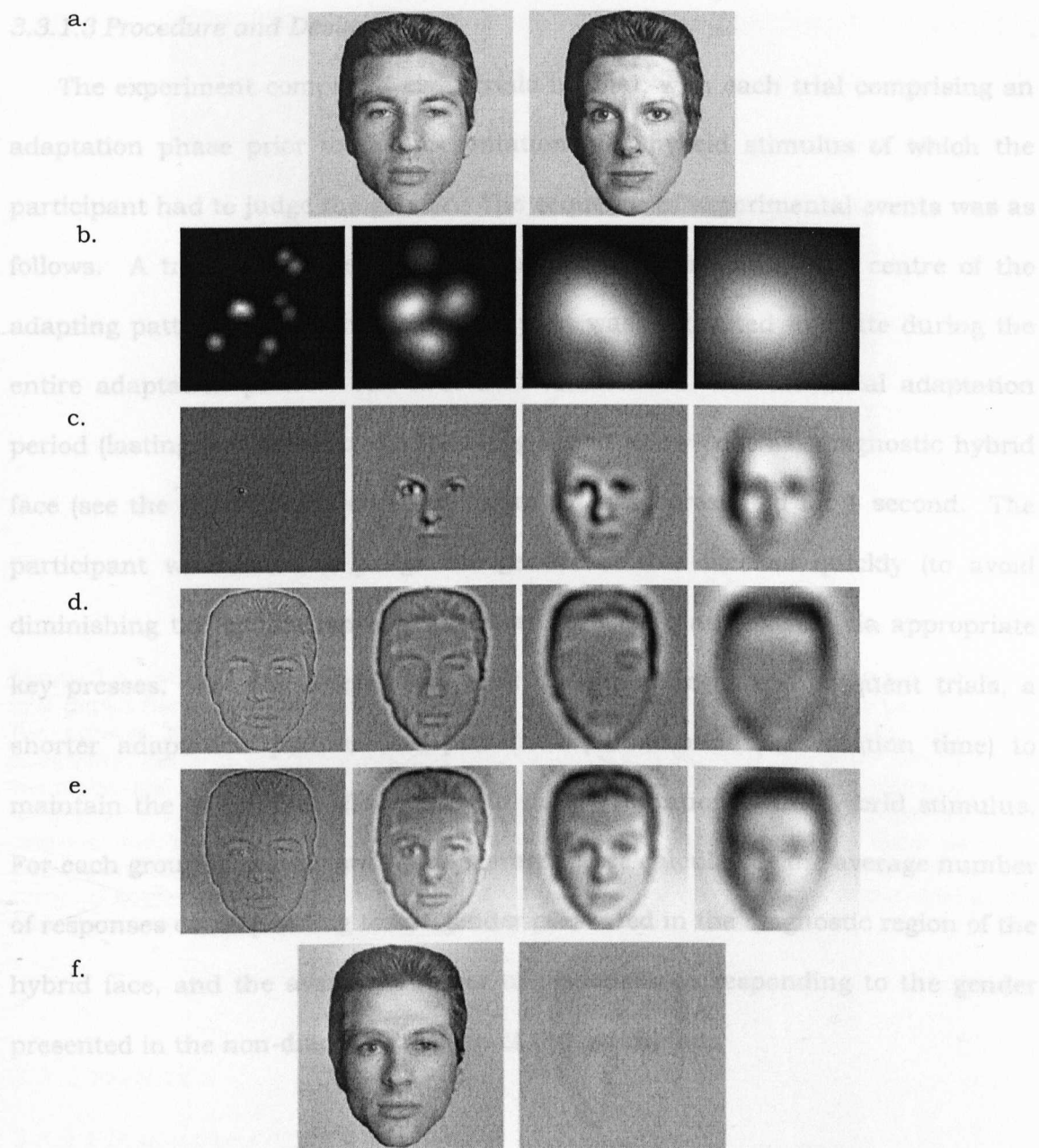


Figure 3.3. Stimulus generation for Experiment 3.2. The pictures in (a) represent two of the original faces used to generate the hybrids, a male on the left and female on the right; (b) illustrates the diagnostic masks for the gender task, left to right, from Schyns et al. (2002), from fine to coarse. The white regions delineate the diagnostic regions and the black the non-diagnostic regions. In (c) the female face in (a) is filtered into 5 different bandwidths (only 4 are shown here as the fifth did not contain any useful information) and is multiplied by the filters in (b) to reveal the diagnostic information of the face; the same operation is carried out in (d), the male face is inserted into the non-diagnostic regions. The pictures in (e) show the addition of (c) and (d) at each scale. A reconstruction of the images in (e) is shown in the left-hand picture of (f) and the right-hand picture is a low-contrast version presented to observers.

3.3.1.3 Procedure and Design

The experiment comprised eight trials in total, with each trial comprising an adaptation phase prior to the presentation of a hybrid stimulus of which the participant had to judge the gender. The sequence of experimental events was as follows. A trial commenced with a fixation cross, located in the centre of the adapting pattern, on which each participant was instructed to fixate during the entire adaptation phase. The first trial commenced with an initial adaptation period (lasting for 360 seconds) following which a low-contrast diagnostic hybrid face (see the right-hand picture in Figure 3.4) was presented for 1 second. The participant was asked to judge the gender of this face as quickly (to avoid diminishing the adaptation effect) and as accurately as possible via appropriate key presses. For the second trial, and the remaining six subsequent trials, a shorter adaptation pattern was presented (30 seconds presentation time) to maintain the adaptation effect, prior to the presentation of the hybrid stimulus. For each group of participants, two averages were calculated, the average number of responses corresponding to the gender presented in the diagnostic region of the hybrid face, and the average number of responses corresponding to the gender presented in the non-diagnostic region of the hybrid face.

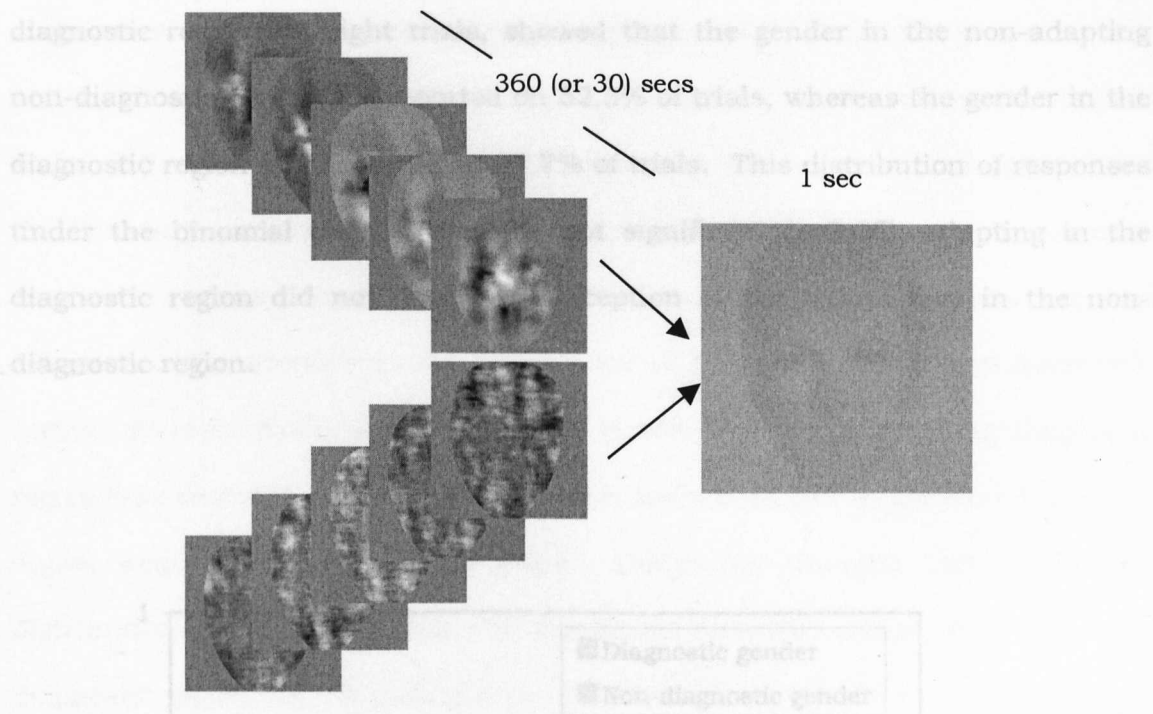


Figure 3.4. Stimuli and design of Experiment 3.2. The smaller pictures on the left-hand side depict the adaptation stimuli, generated by colouring 1000 white noise patterns with the average spatial frequencies of our stimulus set. The top pictures depict the adaptation patterns for the ‘diagnostic-adaptation’ condition with the noise inserted into the diagnostic regions, at each spatial scale, cropped by a face-mask. The pictures in the bottom left-hand side illustrate the ‘non-diagnostic adaptation’ condition with the noise inserted into the non-diagnostic regions at each scale. Two hundred of these patterns were repeatedly presented dynamically in a movie. The right most picture is a low-contrast diagnostic hybrid comprising a female in the diagnostic regions and a male in the non-diagnostic regions.

3.3.2 Results

3.3.2.1 Diagnostic-adaptation condition

If adapting to the spatial frequency information underlying the average hybrid face in the diagnostic region suppresses the information for accurate gender recognition of the face in this region, then observers should report the gender of the face in the non-diagnostic region. An analysis of the data provided by five observers who adapted to the average spatial frequencies of the faces in the

diagnostic region, on eight trials, showed that the gender in the non-adapting non-diagnostic region was reported on 52.3% of trials, whereas the gender in the diagnostic region was reported on 47.7% of trials. This distribution of responses under the binomial distribution was not significant ($p>0.05$); adapting in the diagnostic region did not induce a perception of the hybrid face in the non-diagnostic region.

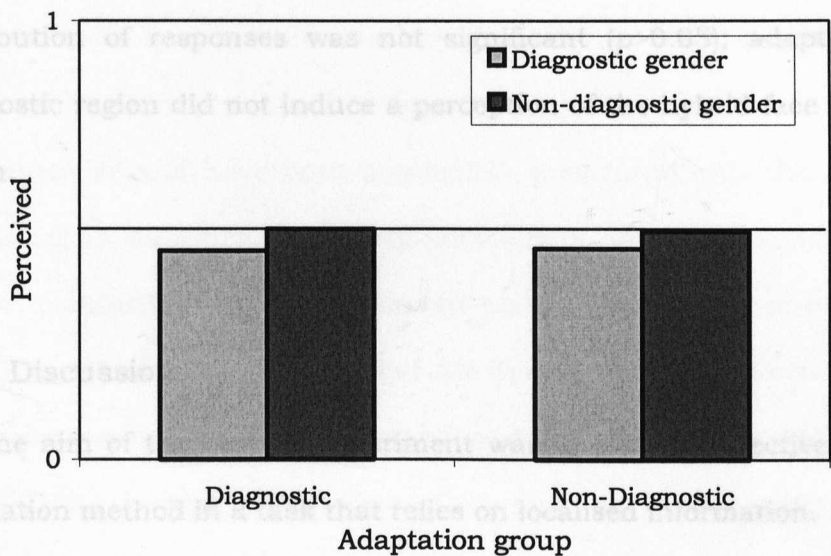


Figure 3.5. The results of Experiment 3.2. The bar graph shows the proportion of responses for each adaptation group.

3.3.2.2 Non-diagnostic adaptation condition

Following the reasoning of the expected result of the Diagnostic adaptation condition we would expect that if adapting in the non-diagnostic region

suppresses the local information necessary to correctly judge the gender in this region, then the observer should report the gender in the diagnostic region. Figure 3.5 shows the frequency of responses, for the Diagnostic-adaptation and the Non-diagnostic adaptation groups, for the gender in the diagnostic and non-diagnostic regions. An analysis of the data provided by thirteen observers who adapted to the average spatial frequencies of the faces in the non-diagnostic region, on eight trials, showed that the gender in the non-adapting diagnostic region was reported on 48% of trials, whereas the gender in the non-diagnostic region was reported on 52% of trials. Under the binomial distribution this distribution of responses was not significant ($p>0.05$); adapting in the non-diagnostic region did not induce a perception of the hybrid face in the diagnostic region.

3.3.3 Discussion

The aim of the current experiment was to test the effectiveness of the local adaptation method in a task that relies on localised information. We attempted to induce a perception of a hybrid stimulus that comprised a face of one gender in the diagnostic regions and a face of another gender in the non-diagnostic regions, that was opposite of the adapting region. Locally adapting to the spatial frequencies underlying the average hybrid face did not significantly result in a categorisation of the face in the non-adapting region. There are a number of potential explanations for these null results. Firstly, it is possible that the adaptation phase failed the challenge of adapting simultaneously the many spatial frequencies at multiple orientations, all of which had to be adapted to

obtain the desired effect. This argument however is weakened by the successful adaptations observed in Experiments 2.2 and 3.1. Perhaps more relevant is the possibility that the hybrid stimuli ineffectively portrayed two genders. The imbalance in the spatial extent of the diagnostic and non-diagnostic regions could have resulted in an inherent bias towards perceiving the gender represented in the predominant non-diagnostic region. Although we attempted to transcend this bias by differentially adjusting the contrast of the faces in each region, the results here do not greatly deviate from pilot work where the hybrid stimuli retained uniform contrast. Perhaps a more fundamental strategy to deal with this problem would have been to establish the difficulty of the gender recognition task of hybrid faces prior to conducting the adaptation experiment. A preliminary experiment should have been conducted, presenting only the hybrid faces and subsequently recording gender categorisations, with no adaptation phase. Our failure to include such an experiment testing the effectiveness of these stimuli, therefore, raises the possibility that the hybrid stimuli ineffectively portrayed two alternative percepts. A further methodological concern is the lack of reaction time data for the task; despite being instructed to respond as fast as possible, the responses of observers may have been too long, resulting in a less strong adaptation effect.

In considering the aim and outcome of the current experiment, perhaps an issue of greater interest relates to the conception of diagnostic information, as presented here. Here, we used the gender masks from Schyns et al. (2002) generated from a bubbles experiment, to assign a feature's diagnostic status. In that experiment a feature was deemed diagnostic if it reached a specified threshold. Yet, in a practical sense, the level at which this threshold is set is

entirely subjective according mainly with the whim of the experimenter. The assignment of a feature as diagnostic, or non-diagnostic, for a particular visual task is therefore problematic. It is more likely the case that features follow a gradient, rather than a definitive, diagnostic status. The distinction, therefore, between the sufficiency, preference or necessity of a feature in resolving a perceptual task, is blurred. An additional and related point regards the strategies employed by individual participants in resolving a recognition task due to the possibility that there are gross individual differences in the facial regions observers use in resolving such a task. The gender masks used here, derived from the bubbles experiment of Schyns et al. (2002), represent the average regions across observers, an experiment using the bubbles method could easily investigate individual differences in the information observers use to resolve face categorisation tasks.

3.3.3 Concluding remarks

Using local frequency-specific adaptation in either the diagnostic regions of a face, or non-diagnostic regions, our goal was to induce a gender decision of a hybrid face that reflected the face in the non-adapting regions. Due perhaps to the inherent bias in the hybrid stimuli themselves in terms of the imbalance in information, the desired effect was not obtained. Further experimentation should focus on establishing the effectiveness of the hybrid stimuli in representing two percepts prior to testing the effect of local adaptation on subsequent gender judgements.

Chapter 4

Learning to Discriminate Features in Noise¹⁰

An abundance of data has been collected over the years, demonstrating improvements in performance on several perceptual and recognition tasks, as a result of learning and attention, and over time, practice in detection or discrimination tasks, can lead to performance improvements that are long lasting, are specific to the experimental conditions under observation, and occur in the absence of feedback (for a review, see Fine and Jacobs, 2002). These phenomena are generally associated with perceptual learning.

Outside the laboratory, real world examples of perceptual learning are widely evident. Consider the ability of the wine or whisky expert who can make the fine discriminations necessary to notice the difference between two wines of a different vintage from the same vineyard, or to detect the ‘nose’ of quality single malts. In the visual sense, medical experts develop the ability to discriminate normal from abnormal structures in X-ray images, which to the novice appear like random

¹⁰ Experiment 4.1, described in this chapter, was presented as a poster presentation at the European Conference on Visual Perception (ECVP), 2002. Bonnar L., & Schyns, P.G. (2002). Finding Diagnostic Features in Noise. *Perception*, **31** (Supp), 18.

blobs. In recent years, methodological developments have led to considerable research activity addressing the mechanisms behind performance improvements on a range of perceptual tasks, including those that lead to perceptual learning.

4.1 Mechanisms of perceptual learning

In discrimination tasks that rely on making fine distinctions between different instances of the same class of stimuli, such as a face identification task, an efficient strategy, considering resource limitations, would be to attend to the information that is optimal to resolve the task, that is, the diagnostic information. In order to successfully extract this useful information in the context of other competing information, the visual system must firstly learn *what* the information is so that attention can be deployed to extract it.

The potential mechanisms underlying improvements in performance on perceptual tasks involving selective attention and perceptual learning were introduced in Section 1.4.2. To recap, a mechanism of signal enhancement operates to augment the strength of the incoming signal, a mechanism of external noise exclusion narrowly tunes the perceptual template to exclude any external noise, and lastly, a mechanism of internal noise reduction reduces the internal neural noise naturally associated with perceptual processing. The roles of these mechanisms in underlying the performance improvements accompanied by attention and perceptual learning have been investigated using the methods of noise masking and response consistency.

Using the method of noise masking, Lu and Doshier (2004) distinguished between these mechanisms to investigate the mechanism(s) behind improvements

in performance in an orientation discrimination task, as a result of perceptual learning. To distinguish between these mechanisms, the contrast energy of the signal was varied and different levels of external noise were added to the signal. Each of these mechanisms carries different predictions regarding performance in these experimental conditions; different performance signatures are revealed when the contrast necessary for a pre-specified performance criterion is plotted against each level of external noise. In this task, it was found that the mechanism behind performance improvements in an orientation discrimination task in the fovea was external noise reduction.

The method has also been fruitfully applied to distinguish between the mechanisms behind the perceptual learning of faces and textures (Gold et al., 1999b; 2004). The mechanism behind the observed improvements (up to 400% discrimination improvement over several days) in face and texture discrimination tasks was signal enhancement. Comparisons with an ideal observer, that was able to use all of the available information in the stimulus to resolve the task, revealed that the human observer, over time, became more efficient at encoding relevant information for accurate discrimination. To my knowledge, this study was the first to apply a noise masking paradigm to high-level stimuli such as faces, to reveal the mechanism that operates to support learning of such stimuli. However, since the contrast energy of the signal and the external noise added to the signal were uniform across the stimulus, the results of Gold et al. (1999b) do not inform us of the locality of this learning. A later study by Gold et al. (2004) employed the method of reverse correlation to trace the changes in the facial regions observers used to discriminate two faces and two texture patterns. Focussing on the results from their face discrimination task, the classification

images of Gold et al. on the first half of learning, show a bias towards using information from the left side of the face, in particular the left eye region, then in the second half of learning, a greater use of the right eye region emerges. Their results also showed an increase in the similarity between their participants' template in resolving the 2-AFC face discrimination task, and that of an ideal observer that can use all of the available stimulus information.

The study of Gold et al. goes a long way in identifying both the mechanism underlying the perceptual learning of faces and in attempting to trace this learning to particular facial regions (their Experiments 1 and 3, respectively). But does the suggested mechanism of signal enhancement operate across the entire image? In other words, is it possible that a mechanism of signal enhancement operates to improve the efficiency with which the observer samples information from diagnostic regions of the face, and another mechanism functions to exclude less useful facial regions? To obtain an online map of the perceptual learning mechanisms in operation in this task, across different facial regions, a method is required that simultaneously informs us of the mechanism(s) in operation, and the facial regions over which they operate. Here, I apply the noise masking method in a perceptual learning paradigm involving the discrimination of ten faces to map learning through time to specific face regions.

4.2 Experiment 4.1

The experiment uses the method of noise masking, in a perceptual learning paradigm, to examine the mechanism(s) behind any discrimination improvements over time and to localise these improvements to particular features. Previous

work, using the *Bubbles* method, has revealed the features used to resolve an identity task (Gosselin and Schyns, 2001; Schyns et al., 2002). In these experiments, observers used, across different spatial scales, the eyes, the mouth and the chin. To assess the locality of any learning effects, and to obtain an online map of the attentional mechanisms in operation in this task, we applied regional masks that revealed information at different regions of the face (eg. the left eye). The contrast energy of the signal, revealed by each regional mask, was differentially adjusted, and different levels of external noise were added to each region, to map learning onto specific facial features and the mechanism(s) underlying this learning.

4.2.1 Methods

4.2.1.1 Participants

Three consenting University of Glasgow students with normal, or corrected to normal vision were paid for their participation in this experiment.

4.2.1.2 Stimuli

The experiment ran on a Macintosh G4 using a programme written with the Psychophysics Toolbox for Matlab (Brainard, 1997; Pelli, 1997). Stimuli were generated using the greyscale faces of Schyns and Oliva (1999) (five males and five females with normalised hairstyles, orientation and lighting, centred on a 256 x 256 grey background of average luminance). Figure 4.1 illustrates the stimulus generation process. To assess learning at a region-by-region level, using Photoshop® to generate six regional face masks (the regions were left eye, right

eye, nose, left mouth, right mouth, and facial outline) that when applied to each face, revealed information in each respective region. To remove any fine edges, each mask was low-passed using a Gaussian filter (radius of 5 pixels). The application of these masks to the face stimulus allowed the contrast and noise to be independently adjusted at each region. To vary the contrast level of the signal and the density of the noise added to the signal, we used the method of constant stimuli. We used five levels of contrast and five levels of noise density, identified by pilot studies to span the threshold range. The stimuli, observers actually saw, was generated in the following way.

Prior to each trial, a pseudo-random number generator in MatLab selected the face to be presented, and for each region of the face, one of five contrast and noise levels. The adjustment in contrast and the addition of noise was carried out separately for each region. I began firstly by adjusting the contrast of the face. Next, to the reduced contrast face, I added, pixel-by-pixel, a randomly generated Gaussian noise pattern (256 x 256 pixels to match the size of the face image) with a mean of zero, and a variance that was 1 of 5 δ values (0.05, 0.087, 0.15, 0.259, 0.45), each indicating, from low-to-high, the density of noise. The regional mask was then applied to the face to crop the signal to this region. This procedure was followed for each of the six facial regions: left eye, right eye, nose, left mouth, right mouth, and face shape. Finally, I added together each region to create a face that varied, region-by-region, in contrast and the density of added external noise. Following the first day of training, contrast levels were established daily for each observer to avoid ceiling performance. The stimulus was centred on a 256 x 256 background of average luminance and subtended 5.72 x 5.72 deg of visual angle.

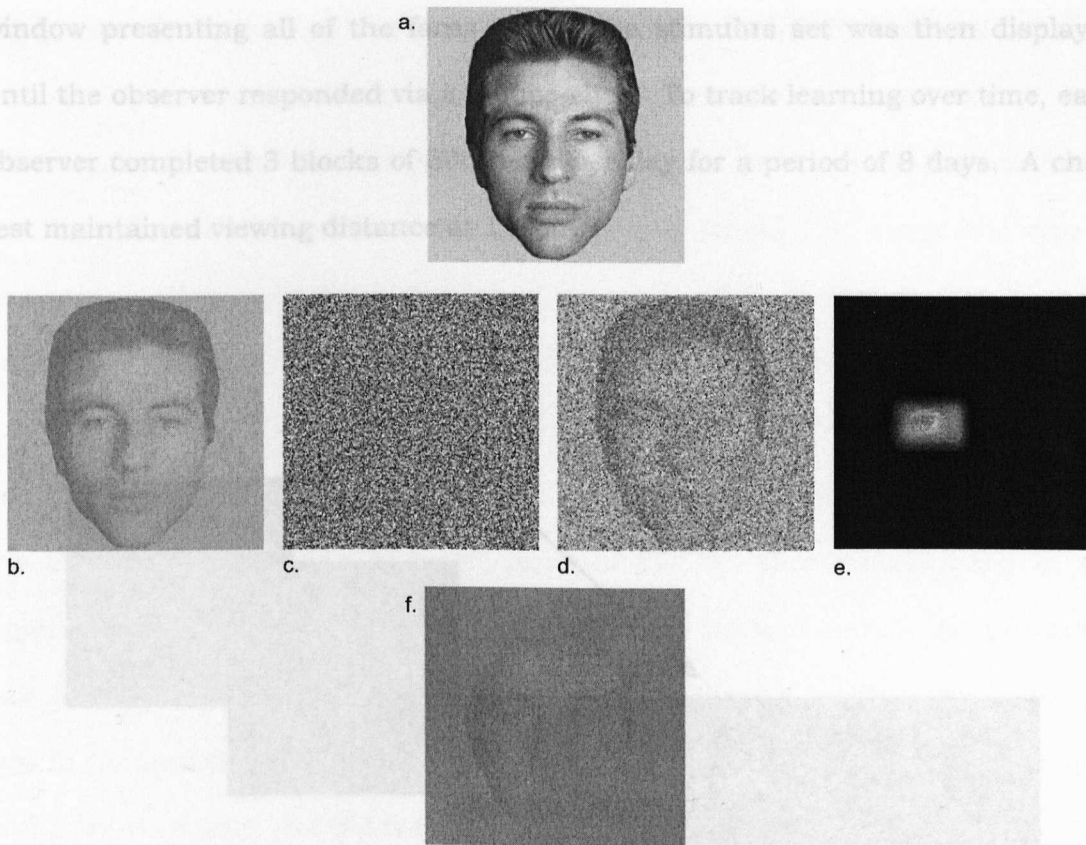


Figure 4.1. The stimulus generation process. The pictures in (a) through to (d) illustrate the stimulus generation process for one region (here, the left eye), but please note that the same procedure is followed on the same original face for different contrast energy levels and noise density levels in each of the six regions. One of the original faces from our stimulus set is depicted in (a); (b) shows the reduced contrast version of (a) and (c) shows a randomly generated white Gaussian noise pattern that is added to (b) to obtain (d). The image in (e) shows the application of one of the region masks (here, the left forehead) to the image in (d). The final image in (f) shows an example of the stimuli observers were presented; it is the result of the addition of steps (a) to (e) for each of the six regions. It depicts a face with varying levels of contrast energy and different levels of noise densities.

4.2.1.3 Procedure

4.2.1.3.1 Human Observers

The sequence of events for a single trial is depicted in Figure 4.2. Each trial began with a fixation point, presented in the centre of the screen for 1 second, following which the test stimulus was presented for 500ms. A response selection

window presenting all of the faces within the stimulus set was then displayed until the observer responded via a mouse-click. To track learning over time, each observer completed 3 blocks of 500 trials per day for a period of 8 days. A chin-rest maintained viewing distance at 100cm.

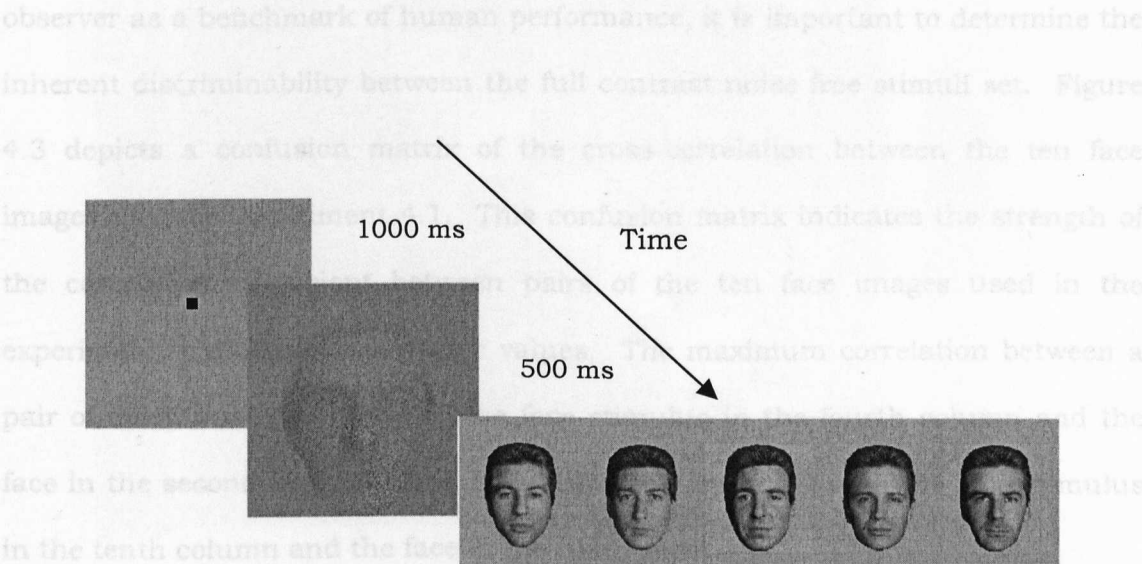


Figure 4.2. Example of a single trial in Experiment 4.1. A trial commenced with a fixation point on a blank field for 1000 ms. A target stimulus, a face was then presented for 500 ms, followed by a response window containing thumbnail-sized pictures of each of the ten face images. To respond, the observer had to select one of these images via a mouse click.

4.2.1.3.2 Ideal Observer

An ideal observer, performing an identical task to the human observer, was presented with the same face stimuli as the human observer. This theoretical observer takes into account all of the information available, pixel-by-pixel, in the signal. The observer uses a decision rule whereby the cross correlation between

the presented signal and each of the full contrast noise-free templates it has in memory is maximised to resolve the 1-of-10 face discrimination task. In a winner-take-all fashion, the observer simply chooses the template that results in the highest cross correlation with the presented signal. In using this type of observer as a benchmark of human performance, it is important to determine the inherent discriminability between the full contrast noise free stimuli set. Figure 4.3 depicts a confusion matrix of the cross-correlation between the ten face images used in Experiment 4.1. This confusion matrix indicates the strength of the correlation coefficient between pairs of the ten face images used in the experiment, indicated in greyscale values. The maximum correlation between a pair of faces was .929 (between the face stimulus in the fourth column and the face in the second row), and the minimum .756 (here, between the face stimulus in the tenth column and the face in the ninth row).

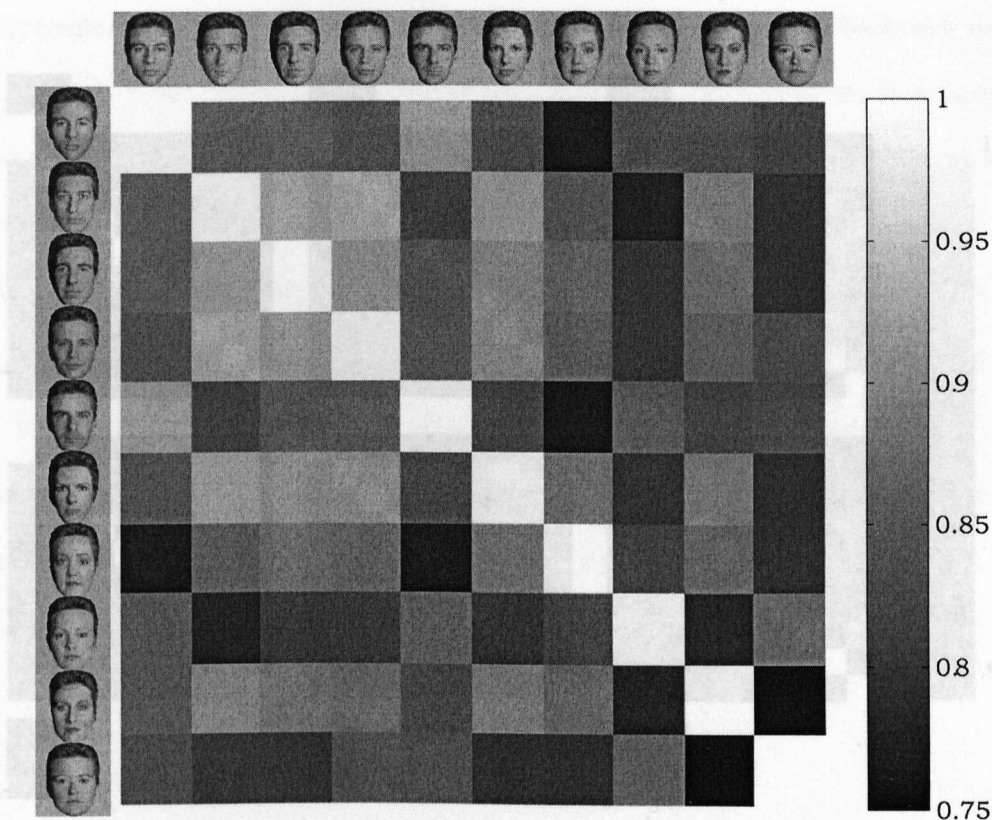


Figure 4.3. Confusion matrices for each of the different face regions, derived from the ten faces used in Experiment 4.1. The maximum and minimum correlations between two

Figure 4.3. The confusion matrix for the image set used in Experiment 4.1. The figure depicts the cross-correlation, in greyscale, between each of the ten face images used in Experiment 4.1; along the columns at the top are faces 1 to 10, left to right, and down the rows from top to bottom. The colour bar indicates the strength of the correlation; the minimum correlation between two of the ten faces was 0.756 and the maximum 0.929.

4.2.1.4 Discrimination

We also computed a confusion matrix, depicted in Figure 4.4, for each of the six facial regions, the left eye, right eye, nose, left mouth, right mouth, and shape. In the left eye region of the ten faces, the maximum correlation was .966 and the minimum, .896; .981 and .957 for the right eye; .991 and .941 for the nose region; for the left mouth .986 and .921; .997 and .982 for the right mouth; and .987 and .930 for the shape.

4.2.1.4 Discrimination

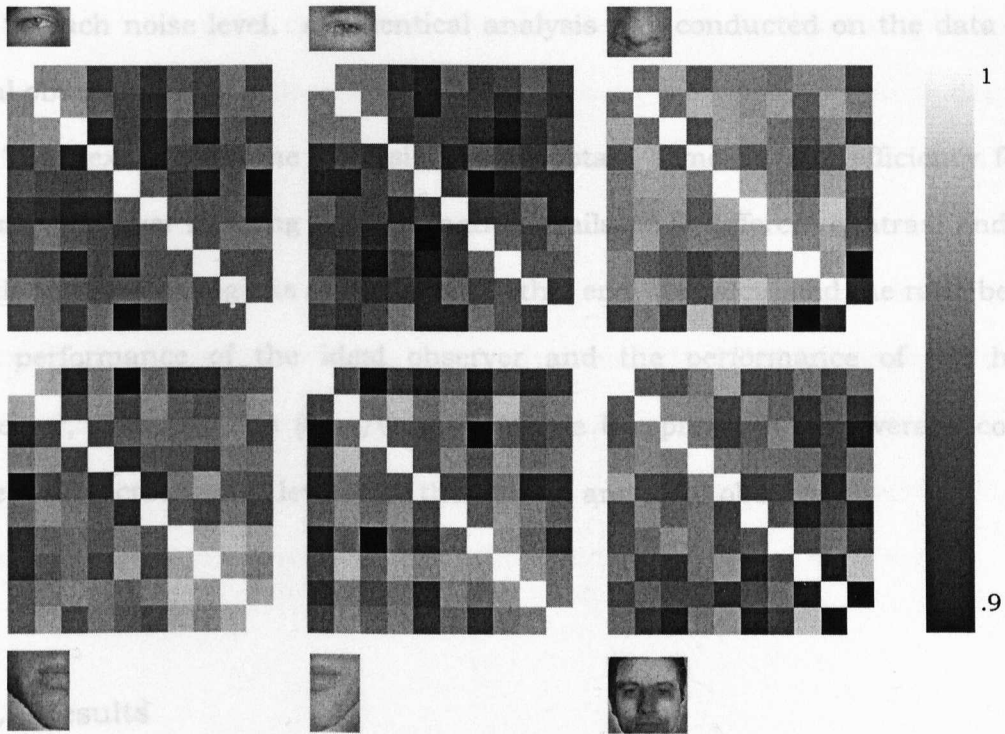


Figure 4.4. Confusion matrices for each of the different face regions, derived from the ten faces used in Experiment 4.1. The maximum and minimum correlations between two image features of the ten face images used were as follows (column and row co-ordinates in parentheses): left eye .966 (2,3) and .896 (5,10), maximum and minimum respectively; right eye .981 (2,1) and .957 (5,7); nose .991 (6,9) and .941 (10,8); left mouth .986 (7,9) and .921 (1,5); right mouth .997 (2,4) and .982 (7,3); and lastly, for shape the maximum and minimum correlations were .987 (8,10) and .930 (10,9), respectively.

4.2.1.4 Data Analysis

For each observer and learning session, responses were recorded in matrices each corresponding to the five contrasts and noise levels for each of the six regions of the face. To determine a contrast threshold at a specific performance criterion for each noise density level at each facial region and learning session, a cumulative Gaussian was fitted to the mean discrimination performance of each cell in these matrices, yielding a discrimination contrast threshold at a

performance criterion of 60% for Observer D.K. and 65% for Observers S.H. and V.L. at each noise level. An identical analysis was conducted on the data of the ideal observer.

The next step in the analysis was to obtain a measure of efficiency for the human observer in using the information available at different contrast and noise levels in different regions of the face. To this end, we calculated the ratio between the performance of the ideal observer and the performance of the human observer, given by $F = (\theta_{\text{ideal}}/\theta_{\text{Human}})^2$, where θ represents the average contrast threshold (across noise levels) for the human and ideal observers.

4.2.2 Results

In order to gain stable estimates of contrast thresholds, the data for each observer was collapsed over the following learning sessions. For Observers D.K. and S.H., data from eight days of learning were collapsed into two sessions: Session 1 (Days 1, 2, 3, and 4) and Session 2 (Days 5, 6, 7, and 8). For Observer V.L., data were collapsed into four learning sessions: Session 1 (Days 1 and 2), Session 2 (Days 3 and 4), Session 3 (Days 5 and 6), and Session 4 (Days 7 and 8).

Figure 4.5 depicts the outcome of the fitting procedure used to derive contrast thresholds at each level of noise density for each region of the face, for Observer D.K. in the left eye region of the face for Session 1 in (a) and Session 2 in (b). The figure depicts in (a) and (b) panels plotting the proportion correct for five different contrast levels (here, contrast is measured on a log scale) at five levels of external noise. Note that contrast increases in intensity from left-to-right within each plot, and in noise density from the top left-most plot to the bottom right-most plot in

(a). The horizontal line in each plot marks a performance criterion of 60%. Note, albeit with some exceptions, that proportion correct increases as contrast energy increases, and decreases as noise density increases. In some cases, the fitting procedure yielded a threshold that was outside the range used in the experiment, resulting in no threshold being recorded. Consistently, across all three observers, this occurred in calculating thresholds for the nose region; this region is therefore omitted from any further analysis. This fitting procedure was followed to obtain contrast thresholds for each of three observers in five levels of noise density in six regions of the face, for each learning session. Identical analyses were also followed to obtain thresholds for the ideal observer.

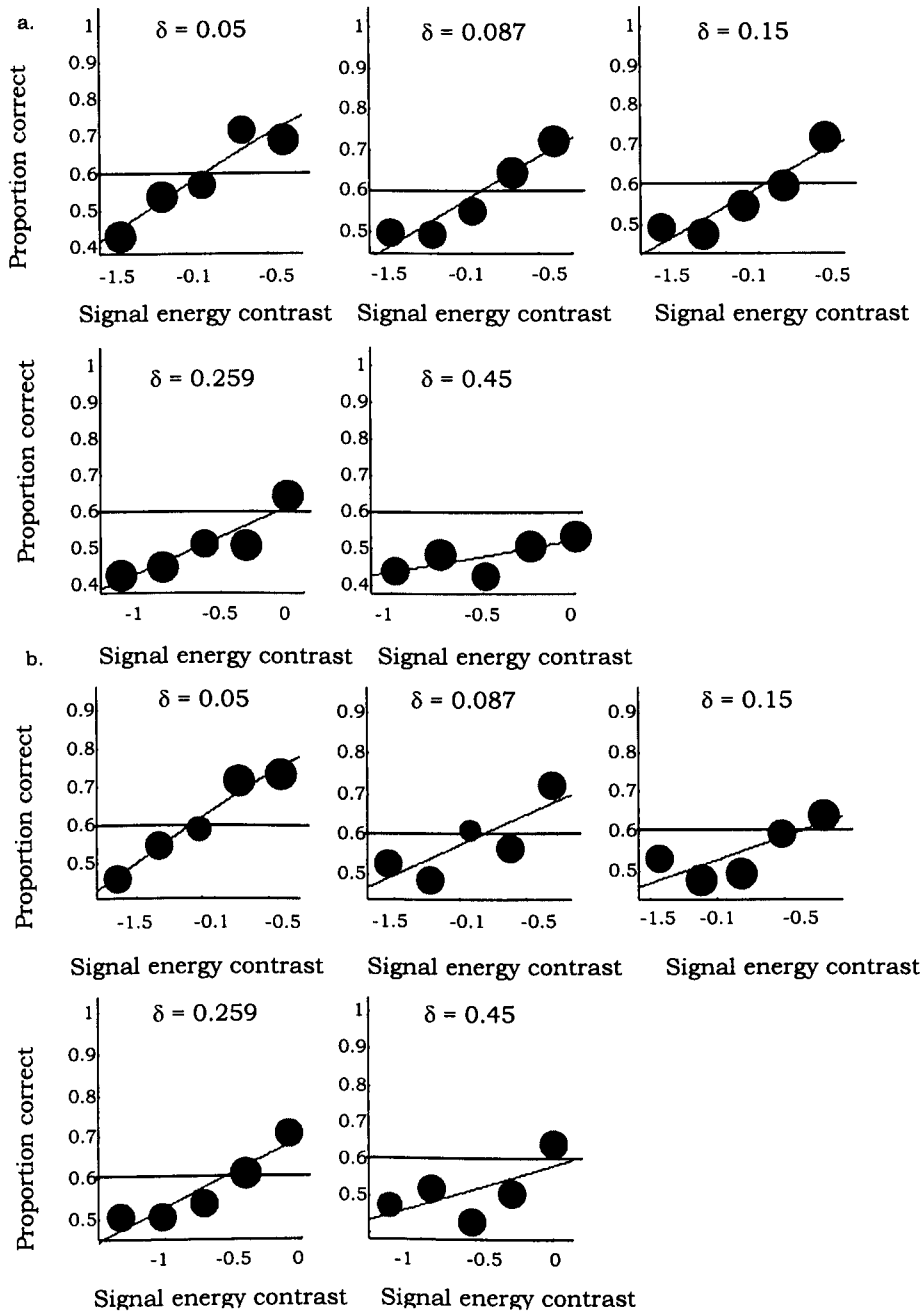


Figure 4.5. Cumulative Gaussian functions plotting proportion correct as a function of signal energy contrast (plotted on a log scale) for Observer D.K. in the face discrimination task of Experiment 4.1. Each of the plots corresponds to performance at a particular level of external noise in one of the six face regions, here the left eye. The five plots in a) correspond to the first two days of learning, and in b) the last two days of learning; for each of the five levels of noise density, increasing from left-to right, top-to-bottom, proportion correct is depicted for the five different levels of contrast tested. These functions were determined for each observer and day of training independently each of the six regions of the face. (The size of each dot relates to the number of times this level of noise and contrast were presented so that a larger dot entails a greater number of presentations).

Tables 4.1, 4.2, and 4.3 show the contrast thresholds for each observer in each facial region at each level of external noise density (empty cells correspond to missing contrast thresholds). Table 4.1 presents the contrast thresholds obtained for Observer D.K. for separate learning sessions: Session 1 and Session 2. With but one exception (the fourth noise density level in the left eye region), the table shows that contrast thresholds increased as a function of external noise density for each of the five regions analysed. Additionally, contrast thresholds decreased from Session 1 to Session 2 (at the third noise level in the left eye region, contrast threshold increased from Session 1 to Session 2).

Table 4.1. Contrast thresholds for Observer D.K. The table depicts contrast thresholds (presented here in % Michelson contrast where performance criterion is at 60% accurate) measured in five different levels of noise density (indicated by δ , as δ increases the density of noise increases) for five facial regions. Thresholds for learning session Days 1 & 2 are shown in grey font, Days 3 & 4 in black font. (Empty cells correspond to missing values).

Facial Region	External noise density				
	$\delta = 0.05$	$\delta = 0.087$	$\delta = 0.15$	$\delta = 0.259$	$\delta = 0.45$
Left eye	0.104	0.172	0.255	0.656	
	0.077	0.139	0.365	0.245	
Right eye	0.133	0.164	0.287	0.834	
	0.082	0.147	0.172	0.465	
Left mouth	0.124	0.180	0.370		
	0.07	0.111	0.271		
Right mouth	0.131	0.259	0.613		
	0.075	0.193	0.201		
Shape	0.109	0.137	0.30	0.674	
	0.076	0.08	0.19	0.360	0.540

The results for Observer S.H. are shown in Table 4.2. At a performance criterion of 65%, contrast thresholds increased as external noise density

increased in all facial regions (with the exception of a minor decrease in threshold as external noise increased from the lowest level to the next level of density in the left eye region). Importantly, the effect of learning on reducing contrast thresholds was evident in the decreases between Session 1 and Session 2 for all facial regions.

Table 4.2. Contrast thresholds for Observer S.H. The table shows contrast thresholds (% Michelson contrast where performance criterion is at 65% accurate) measured in five different levels of noise density for five facial regions. Thresholds for Session 1 are shown in grey font, Session 2 in black font. (Empty cells correspond to missing values).

Facial Region	External noise density				
	$\delta = 0.05$	$\delta = 0.087$	$\delta = 0.15$	$\delta = 0.259$	$\delta = 0.45$
Left eye	0.14	0.324	0.659		
	0.03	0.02	0.04	0.042	0.044
Right eye	0.128	0.169	0.372		
	0.016		0.062	0.03	0.036
Left mouth	0.114	0.17	0.428	0.799	
	0.019	0.021	0.041	0.032	0.067
Right mouth	0.181	0.214	0.345	0.789	
	0.015	0.029	0.038	0.03	0.036
Shape	0.144	0.164	0.257	0.851	0.853
	0.013	0.019	0.032	0.065	0.192

The contrast thresholds for Observer V.L. at a performance criterion of 65% correct are shown below in Table 4.3 for four learning sessions. In comparison to Observers D.K. and S.H., the contrast thresholds for Observer V.L. do not consistently increase as external noise density increases. Increases in contrast thresholds, as external noise increases, are evident on Session 1 for the right eye, left mouth, right mouth and shape; on Session 2, for the right eye, left mouth,

right mouth and shape; on Session 3, only for the left eye, left mouth and shape, and not in any region consistently on Session 4.

Table 4.3. Contrast thresholds for Observer V.L. The table shows contrast thresholds (% Michelson contrast) measured in five different levels of noise density for five facial regions (65% performance criterion). Thresholds for Sessions 1 & 3 are shown in grey font, Sessions 2 & 4 in black font. (Empty cells correspond to missing values).

Facial Region	External noise density				
	$\delta = 0.05$	$\delta = 0.087$	$\delta = 0.15$	$\delta = 0.259$	$\delta = 0.45$
Left eye	0.088	0.178	0.711	0.062	0.547
	0.141	0.722	0.09	0.227	0.105
	0.037	0.066	0.092	0.423	
	0.015		0.046	0.015	
Right eye	0.083	0.116	0.166	0.455	
	0.037	0.026	0.137	0.32	
	0.052	0.084	0.045		
	0.011	0.021	0.016	0.053	0.046
Left mouth	0.029	0.094	0.138	0.431	
	0.059	0.091	0.117	0.352	
	0.025	0.052	0.132	0.569	
	0.028	0.016	0.025	0.027	0.091
Right mouth	0.061	0.106	0.513		
	0.052	0.085	0.093	0.176	0.709
	0.051	0.041	0.189	0.155	
	0.013	0.016	0.03	0.025	0.026
Shape	0.042	0.068	0.15	0.23	
	0.056	0.09	0.159	0.204	0.623
	0.026	0.032	0.067	0.168	0.584
	0.022	0.019	0.035	0.083	0.154

Figure 4.6 illustrates the results for Observer V.L.; in each face image the contrast in each region is set to match the corresponding contrast threshold for each region according to the different learning session at a performance criterion of 65%. The rows, from top-to-bottom, depict the face images for different training sessions, Session 1 through to Session 4, whereas each column, from left-to-right, represents a different level of noise density, from low-to-high, respectively.

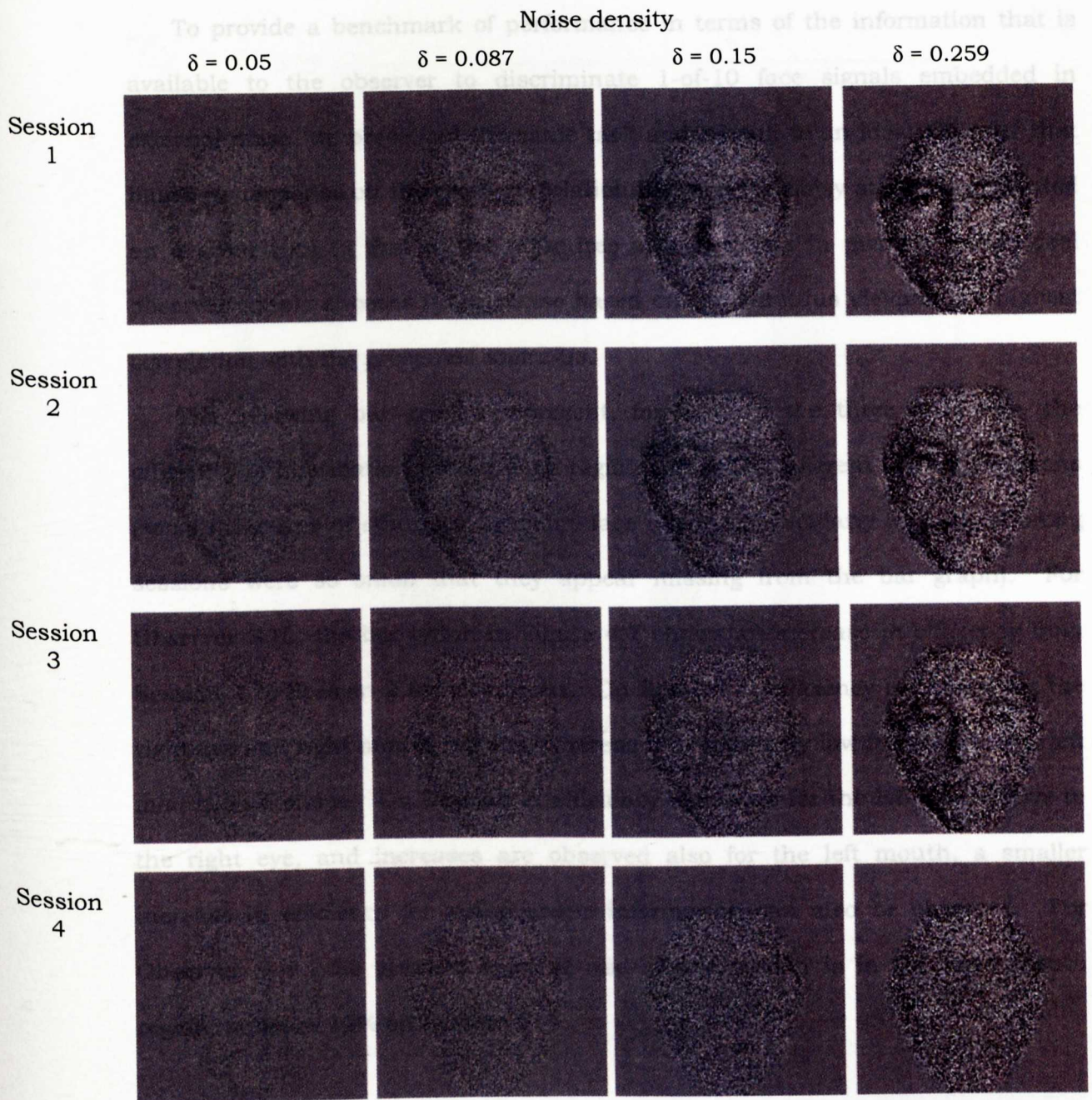


Figure 4.6. The results of Observer V.L. depicted as contrast threshold versus noise image maps. Each picture contains the contrast threshold for each region for a single level of noise density, increasing from the left column to the right-most column. Each row represents a different training session, Session 1 on the top row, Session 4 on the bottom row.

To provide a benchmark of performance in terms of the information that is available to the observer to discriminate 1-of-10 face signals embedded in external noise, we presented the same task and stimuli to an ideal observer that based its response on the cross-correlation between the noisy stimulus presented on a given trial to that of the noise-free stimuli it has in memory. The ideal observer simply chooses its response based on the stimulus yielding the highest correlation with the presented stimulus.

The following bar graphs represent, for each of the three observers, the efficiency of information use for each region across the different learning sessions (some measures of efficiency for some face regions particularly on early training sessions were so small that they appear missing from the bar graph). For Observer S.H., the bar graph in Figure 4.7 shows an increase in efficiency from Session 1 to Session 2 for all regions. On Session 1, efficiency is highest for the right eye and right mouth regions, whereas it is extremely low for the left eye, left mouth, and shape. On Session 2, efficiency increases for the left eye relative to the right eye, and increases are observed also for the left mouth, a smaller increase in efficiency for using shape information can also be observed. For Observer S.H., the greatest efficient use of information is in the right mouth region, at below 12% on Session 2.

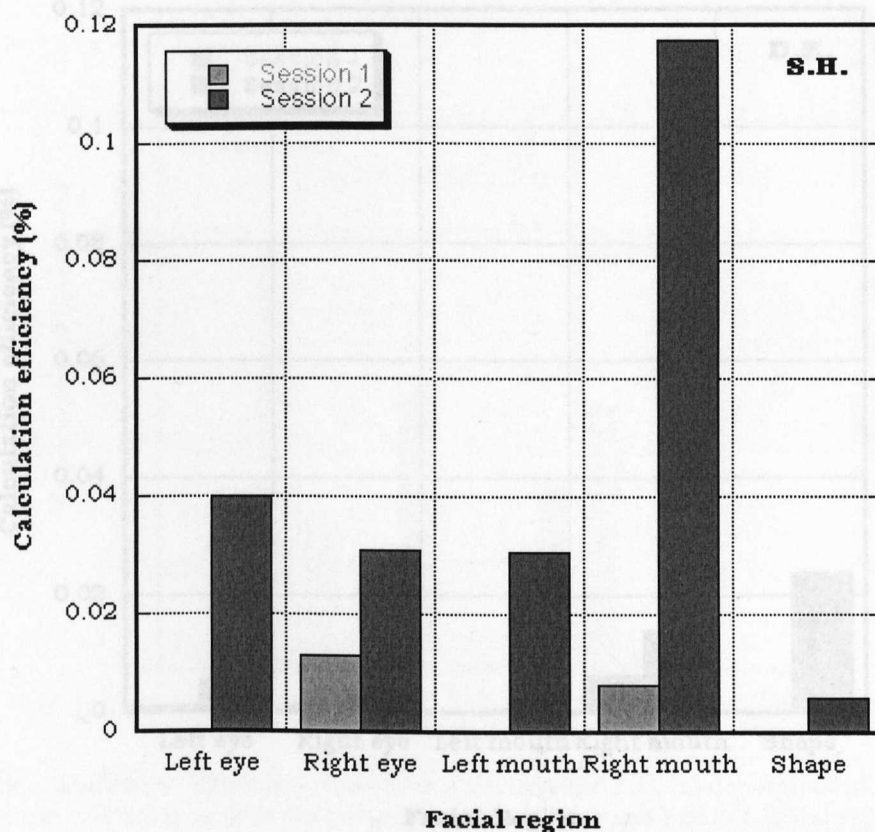


Figure 4.7. Calculation efficiency measures for Observer S.H. The graph depicts measures of calculation efficiency across learning sessions for each of five facial regions.

Figure 4.8 shows calculation efficiency measures for Observer D.K. The bar graph shows lower efficiency levels for Observer D.K. on Session 1 than for Observer S.H. across all regions, with the right mouth showing the highest level of efficiency. On Session 2, efficiency increases for all regions with the exception of the left mouth, and shows the highest increase in using shape information, followed by the right mouth.

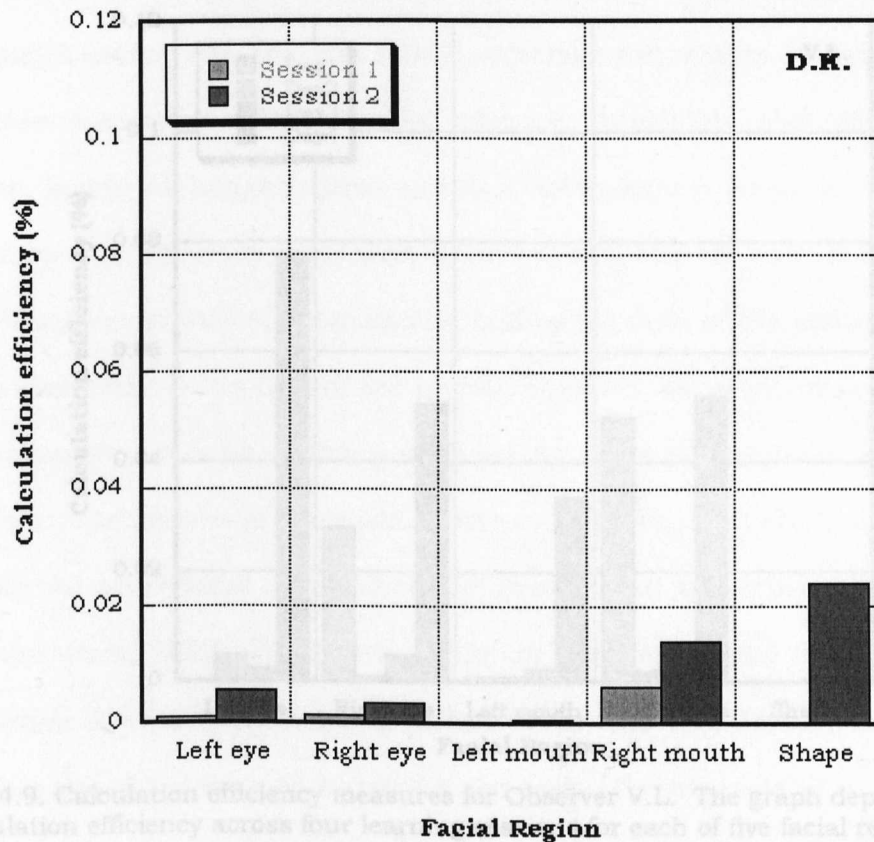


Figure 4.8. Calculation efficiency measures for Observer D.K. The graph depicts measures of calculation efficiency across learning sessions for each of five facial regions.

Lastly, the efficiency measures for Observer V.L., depicted in Figure 4.9, are variable; decreases in efficiency for using information from the left eye, right eye, and right mouth decrease between Session 1 and Session 3, increasing again on Session 4. On Session 1, efficiency is highest for the right eye and right mouth, the left eye on Session 2, the right eye on Session 3, and the left eye on Session 4 at just below 8% efficiency. On Session 4, increases in the efficient use of information are evident for the left eye, right eye, left mouth, and right mouth.

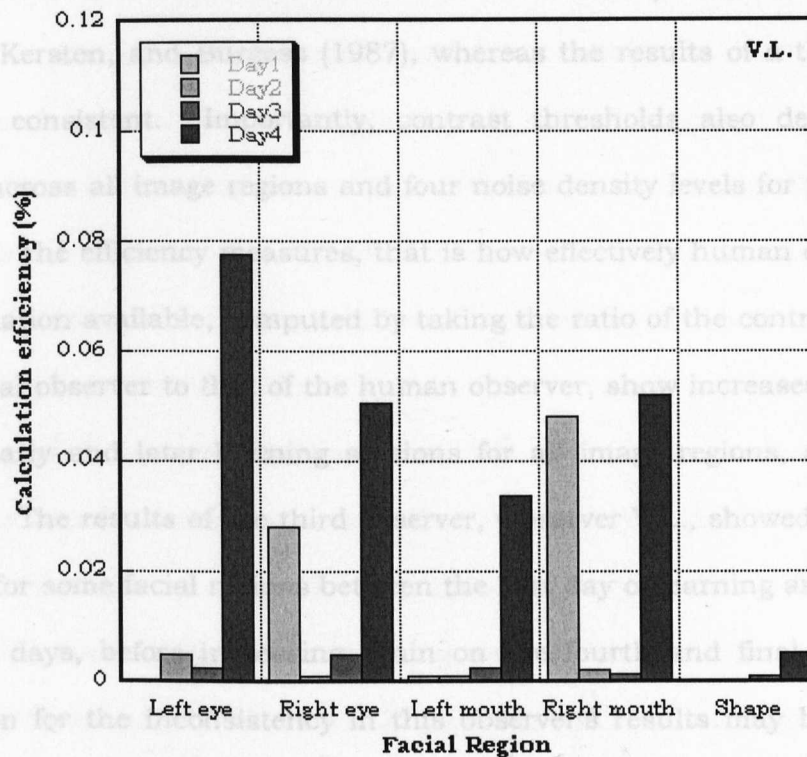


Figure 4.9. Calculation efficiency measures for Observer V.L. The graph depicts measures of calculation efficiency across four learning sessions for each of five facial regions.

Taken together, the results suggest that the efficiency, with which observers use the information available, increases with learning differentially across facial regions, and between observers.

4.2.3 Discussion

In Experiment 4.1, we measured contrast thresholds in varying levels of noise density at different regions of the face over several learning sessions, with the aim of localising learning to specific facial information, and to determine the mechanisms underlying learning. The results of two observers show consistent

increases in contrast thresholds as external noise density increases, as expected by Legge, Kersten, and Burgess (1987), whereas the results of a third observer were less consistent. Importantly, contrast thresholds also decreased with learning, across all image regions and four noise density levels for the same two observers. The efficiency measures, that is how effectively human observers use the information available, computed by taking the ratio of the contrast threshold for the ideal observer to that of the human observer, show increases in efficiency between early and later learning sessions for all image regions, again for two observers. The results of the third observer, Observer V.L., showed decreases in efficiency for some facial regions between the first day of learning and the second and third days, before increasing again on the fourth and final session. An explanation for the inconsistency in this observer's results may be due to the strategies the observer used to resolve the task. For example, it is possible that when the noise was too high in one facial region, the observer simply switched to using the information in another facial region with a lower noise level. That observers may switch automatically to using other less noisy features when the noise level is too high in a preferred feature is an inherent problem with the patchwork nature of the stimulus. Let us consider in more detail the methodological problems with the current experiment.

In regards to localising learning to specific face features, the method we used here treats each feature independently. Yet, it is well documented that observers use configurations of features (for a review see Maurer, Le Grand, and Mondloch, 2002); Schyns et al. (2002) showed that the eyes and mouth are used in conjunction to resolve a face identity task. In the present experiment, we have not examined the role of configural processing in learning to discriminate faces.

It would be possible, however to examine the use of configurations of facial regions in a secondary trial-by-trial analysis, in which the contrast and noise levels in each region are monitored when a correct response is made. The outcome of the current experiment is a face that is contrast modulated across different facial regions for a specified performance criterion and shows how perceptual learning reduces the amount of contrast required across the different facial regions to resolve the task.

The aim of the current experiment was not restricted to localising perceptual learning to specific facial features, rather the aim was also to determine the mechanism(s) enabling learning and subsequent extraction of these image features. We hoped that by manipulating contrast and noise levels differentially across the face image that we could plot, for each facial region, a contrast threshold versus external noise function for each learning session, enabling the mechanism underlying learning to be characterised. For each of the three observers, at some noise levels and face regions, we were unable to establish a contrast threshold that was within the acceptable range (that is within the range used in the experiment), leading to several empty cells. We were unable, therefore, with incomplete data, to fit a model to the data, such as the LAM used by Gold et al. (1999b, 2004) or the PTM of Lu and Doshier (2004), to establish the mechanisms underlying learning. Despite running pilot studies that indicated an initial contrast range, for some regions this range was outside the contrast range required by the observer. Future studies should address this issue by manipulating contrast in a baseline noise-free condition to determine the contrast range per region.

As with all learning experiments, a crucial step in the design is to establish the correct number of trials in a given session in which to assess the effects of learning; an insufficient number of trials may not provide enough observations per condition, whereas too many trials may lead to within-session learning effects. Here, observers completed 1500 trials a day for 8 days; for Observers D.K. and S.H. these trials were split into two learning sessions for subsequent analyses – Session 1 comprising days 1-4, and Session 2 comprising days 5-8, whereas four learning sessions were retained for Observer V.L. Learning was observed between these different learning sessions despite the large number of trials involved, within-session learning effects did not result in a dissipation of cross-session learning effects. However, reducing the number of trials, or the size of the search space itself (for example, measuring thresholds in three levels of noise: low, medium and high density) would enable a closer assessment of any within-session learning effects. This problem of running an extensive number of trials is applicable to other similar methods exploring information use; the reverse correlation method used by Gold et al. in a 2 AFC face discrimination task required 9600 trials from each observer, and the Bubbles method, when used on few observers also requires thousands of trials to sample an entire stimulus space.

A future study aiming to map perceptual learning to specific regions of the face, and to characterise the mechanism(s) enabling learning, could potentially use the Bubbles method of Gosselin and Schyns (2001). Whereas, Gosselin and Schyns searched the face stimulus for diagnostic information using a grey mask randomly punctured with bubbles that revealed local information, one could use instead a gaussian noise mask punctured with bubbles, that from trial-to-trial

varies in noise density, and whose bubbles reveal information at different levels of contrast. Additionally, constraining the search space to the information that has been found to be diagnostic for resolving facial identity (Schyns et al., 2002) would reduce the number of trials required to examine the entire stimulus space.

4.2.4 Concluding remarks

We used the method of noise masking in Experiment 4.1 to map the perceptual learning of a face discrimination task onto specific facial regions, and to characterise the mechanism(s) enabling learning. We observed, for two observers, consistent increases in contrast thresholds as external noise density increased, across facial regions, and a reduction in thresholds with learning. Comparison with an ideal observer that can use all of the available stimulus information provided efficiency measures for each of the different facial regions for each day of learning. Methodological problems, however, that resulted in missing data prevented a determination of the mechanism(s) underlying this learning. We consider the questions of the evolution of information use in perceptual learning, and the corresponding mechanism(s) enabling this learning, as important in achieving both a qualitative, and quantitative description of perceptual learning.

Chapter 5

Attention Modulates the Speed of Visual Information Processing¹¹

A fundamental function of the visual system, and other sensory systems, due to resource limitations, is to optimally select task-relevant information from the barrage of information impinging the retina. The mechanisms that enable such selection are those of visual attention. Visual attention has been shown to enhance performance in various visual tasks, including threshold tasks of contrast sensitivity (Carrasco, Penpeci-Talgar, and Eckstein, 2000), spatial resolution (Yeshurun and Carrasco, 1998, 2000; Carrasco, Williams, and Yeshurun, 2002), and visual search (Carrasco and Yeshurun, 1998). Such evidence supports the view that a core function of visual attention is to facilitate visual processing by enhancing the perceptibility of a range of visual input.

¹¹ The experiments described in this chapter formed a poster presentation at ECVF 2003. Bonnar, L., Jentzsch, I., & Leuthold, H. (2003). Attention and the Speed of Information Processing. *Perception*, **32** (Supp.), 140.

While the effects of attention on the discriminability of visual input have been extensively studied, fewer studies have examined the effects of attention on the temporal dynamics of visual processing. One such study, Carrasco and McElree (2001; see also McElree and Carrasco, 1999), that has examined the effect of attention on visual information processing found acceleration in the rate at which visual information is processed; in other words, attention speeded up processing. If attention facilitates visual processing by altering the temporal dynamics of information processing, at which stage of processing does this modulation occur?

5.1 Determining the locus of a temporal modulation of attention

The locus of any attentional effect within the information-processing stream could potentially lie at an early sensory stage, or at a later, decision-related stage. This chapter seeks to establish the locus of any attentional modulation on the temporal dynamics of processing in a series of three experiments, analysing event-related brain potentials (ERPs), in addition to the accuracy and time course of the behavioural response. To this end, each experiment employed pre-cues to manipulate spatial attention in an orientation discrimination task, and speed-accuracy trade-off (SAT) procedures to examine the full time course of processing. We used the response-signal SAT method that simply consists of presenting an auditory tone, at variable times after target onset, to initiate a limited time window in which the observer must make their response. Using a range of time lags, from short-to-long, between the target and response tone, enables the full time course of processing, and the time window where attention has an effect, to be examined. To establish the processing stages(s) modulated by attention, we

analysed the pre-motoric (S-LRP) and motoric (LRP-R) components of the lateralised readiness potential (LRP). An effect on the pre-motoric S-LRP interval would signify a modulation of early processes, whereas an effect on the motoric LRP-R interval would indicate a modulation of later response-related processes. Previously, Rinkenauer et al. (2004) found that SAT procedures shortened both S-LRP and LRP-R components of the LRP in perceptual and cognitive tasks. We also analysed the effect of attention on the time course and topography of potential amplitude or latency modulations of visual-related ERP components. The earliest ERP components reported modulated by attention, in terms of enhanced amplitude, are the P1 and N1 components (Fu et al., 2001; Hopfinger and Mangun, 1998; Lubbe and Woestenburg, 1997). Yet, in relation to attention speeding up visual processing, no electrophysiological evidence of a shortening of the latency of these early components has been reported, a somewhat surprising situation considering behavioural evidence supporting a modulation on the temporal dynamics of processing. In regards to later components, modulations on the P2, N2, and P3 components have also been observed for valid cues as enhancements in amplitude, and as a shortening of latency of the P3 (Fu et al., 2001), indicating a chronometric effect at the level of stimulus categorisation.

Here, the aim is to establish whether a modulation in the speed of processing expressed on the SAT function may be accompanied by a shortening of the latency in either S-LRP or LRP-R latency, and/or in early ERP components such as the P1/N1, or the later P3. While the use of SAT procedures enable a conjoint investigation of the effects of attention on both the discriminability and temporal dynamics of processing, the localisation of these effects cannot be established with such procedures. The analysis of the LRP and ERPs, however, enable the

locus of such effects to be established. Figure 5.1 illustrates the chain of information processing and the corresponding electrophysiological markers that could potentially be modulated by attention. In relation to the LRP, a chronometric modulation occurring at a pre-motoric stage would be evident as a shortening in the latency of the S-LRP interval, whereas a modulation at the motoric stage would be evident as a shortened latency of the LRP-R interval. In relation to other ERP components, an attentional modulation at sensory stages of processing would be evident on the early P1/N1 components, whereas a chronometric modulation at the later stage of stimulus categorisation would be indexed by an earlier peak latency of the P3 component. Following this stimulus categorisation stage, a later modulation on the ERP waveform could also occur that relates to a response selection stage. The combination, therefore, of employing SAT procedures and recording ERPs, in spatial attention tasks enables both the effect of attention on the temporal dynamics of processing to be characterised, and the locus of this effect to be established.

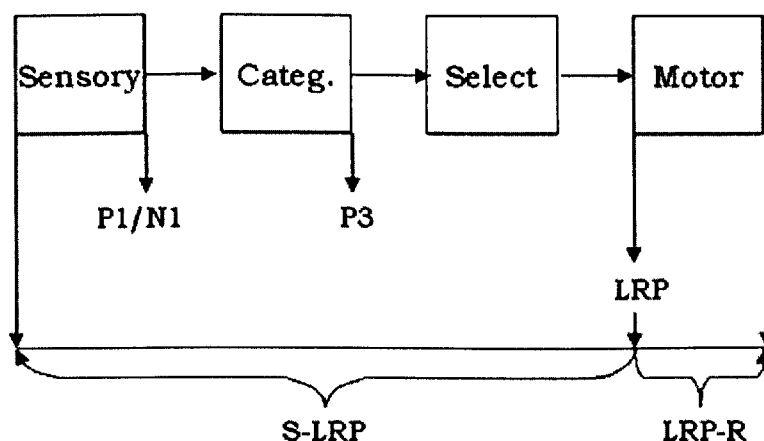


Figure 5.1. The stream of information processing and the electrophysiological markers of processing stages.

In Experiment 5.1, we followed Carrasco and McElree's (2001) methodological and experimental procedures, using a 100% valid peripheral cue that provided information only about target onset and target location, and a neutral cue located centrally providing no information regarding target location. In Experiment 5.2, we used only a peripheral cue but varied its validity ($p(\text{valid})=.5$) to control for the effect of cue presentation (peripheral vs. central, as used in Experiment 5.1) on the ERP waveforms. Experiment 5.3, while retaining the pre-cueing characteristics of Experiment 5.2, employed a more complex display in which the target was always presented amidst distracter stimuli. Here, the observer had to discriminate the target stimulus that appeared at the same location of an adjacent report cue (a small black line) that was either congruent or incongruent with the pre-cue.

5.2 Experiment 5.1

To assess the effect of covert attention on the speed of information processing, we adopted a response-signal speed-accuracy trade-off procedure in a 2-AFC orientation discrimination task, as used previously by Carrasco and McElree (2001). Using an orientation discrimination task in both feature and conjunction search tasks, with and without distracters, Carrasco and McElree found that attention speeded up the rate at which visual information is processed. Using the feature search condition of Carrasco and McElree, in a display without distracters, we manipulated attention using a peripheral valid cue ($p(\text{valid} = 1)$) that indicated the spatial location of the upcoming target, and a central neutral cue that indicated only target onset. As eye movements can be

initiated in less than 150 ms after the abrupt onset of a target, the timing of events was set to preclude eye-movements.

An effect of attention on the discriminability of processing will be revealed as a higher level of asymptotic performance on the SAT function for the cued condition, relative to the neutral condition, whereas an effect on the speed of information processing will be realised on the SAT function as a faster rate of increase in accuracy to asymptotic performance, or as an earlier departure from chance performance. The locus of any effects will be established by analysis of early sensory-related (P1/N1), or later categorisation-related (P3) ERP components. The onset latency of the pre-motoric and motoric intervals of the LRP will also be analysed to determine any chronometric modulation of attention.

5.2.1 Method

5.2.1.1 *Participants*

Participants were twelve consenting students from the University of Glasgow (mean age of 27, 8 females, 11 right-handed) with normal, or corrected to normal vision, and who were paid for their participation.

5.2.1.2 *Stimuli & apparatus*

Participants were tested in a sound-attenuated cabin with no lighting. The experiment ran on an IBM-compatible computer using the Experimental Run Time System (ERTS) software to control the experiment (Dutta, 1995). A response pad was used for response recording with keys operated by the left and right

index finger located about 6 cm to the left and right of the centre. Stimuli were presented on a monitor at a constant viewing distance of 1 m.

A small black square served as a cue stimulus subtending 0.23° of visual angle. In the neutral cue condition it was presented at fixation. In the cued condition, the cue was randomly presented on an invisible compass grid at one of four equidistant locations (north, south, east and west) from the central fixation point. The target stimuli, generated using Matlab, consisted of gabor patches (2 cpd sinusoidal gratings cut by a gaussian of $\delta=X$) that varied only on the orientation dimension, oriented by 12° to the left or right. The contrast of these gabor patches was reduced to 6% (Michelson contrast) based on pilot studies showing correct identification at 75% correct. Each gabor patch subtended 0.8° of visual angle, viewed at a distance of 1m. The target stimulus was randomly presented on an invisible compass grid at one of four equidistant locations (north, south, east and west), 3.89° eccentricity from the central fixation point. The background luminance of the monitor was set to 10.3 cd/m^2 .

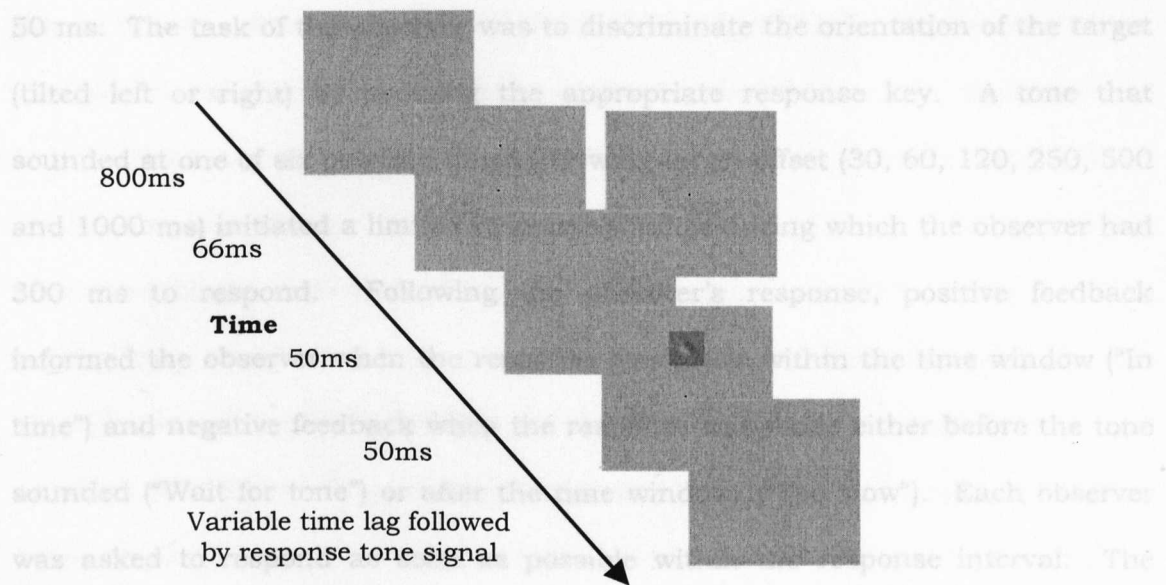


Figure 5.2. A typical trial sequence of Experiment 5.1. Each trial is initiated with a fixation point, presented for 800 ms. A spatial cue, either a peripheral valid cue or neutral central cue, is then presented for 66 ms, followed by an ISI of 50 ms, then the target stimulus, an oriented gabor patch, is presented for 50 ms. After a variable time lag of 30, 60, 120, 250, 500 or 1000 ms, a tone sounds to inform the observer to make their response within a 300 ms time interval. Feedback is then presented ('In Time', 'Too slow' or 'Wait for tone') informing the observer about the timing of their response.

5.2.1.3 Procedure *Physiological Recordings*

Figure 5.2 illustrates a typical trial sequence. Each trial began with a fixation point, presented for 800 ms, on which each participant was asked to fixate throughout the length of the trial. A cue (small black square), peripheral valid or central neutral, was then presented for 66 ms. The neutral cue, which replaced the central fixation point, indicated only the target onset and yielded an equal probability of the target appearing at any one of the four possible target locations. The valid cue was presented adjacent to the upcoming target location, therefore indicating target onset and target location. Following an ISI of 50 ms, the target, an oriented gabor patch, then appeared in one of the four possible locations for

50 ms. The task of the observer was to discriminate the orientation of the target (tilted left or right) by pressing the appropriate response key. A tone that sounded at one of six possible times following target offset (30, 60, 120, 250, 500 and 1000 ms) initiated a limited response window during which the observer had 300 ms to respond. Following the observer's response, positive feedback informed the observer when the response was made within the time window ("In time") and negative feedback when the response was made either before the tone sounded ("Wait for tone") or after the time window ("Too slow"). Each observer was asked to respond as soon as possible within the response interval. The experiment began with a practice block of 43 trials to familiarise participants with the experimental procedure. Each observer then completed 12 blocks of 98 trials in one experimental session (the first 2 trials in each session were practice trials), taking short breaks at the end of each block.

5.2.1.4 Electrophysiological Recordings

Electroencephalographic (EEG) activity was continuously recorded from 70 Ag/AgCl electrodes using a BIOSEMI Active-Two amplifier system, including external electrodes for recording vertical electro-ocular (vEOG) and horizontal EOG (hEOG), at a sampling rate of 256 Hz. EEG activity was recorded from midline electrodes Fpz, Afz, Fz, FCz, Cz, CPz, Pz, POz, Oz, and Iz, over the left hemisphere from electrodes IO1, Fp1, AF3, AF7, F1, F3, F5, F7, F9, FC1, FC3, FC5, FT7, C1, C3, C5, T7, CP1, CP3, CP5, TP7, P1, P3, P5, P7, P9, PO3, PO7, O1, M1, and from the homologue electrodes over the right hemisphere. Figure 5.3 depicts the electrode layout. VEOG and hEOG waveforms were calculated off-

line as follows: $vEOG(t) = FP1(t)$ minus $IO1(t)$, and $hEOG(t) = F9(t)$ minus $F10(t)$. Two additional electrodes, Common Mode Sense (CMS) active electrode and Driven Right Leg (DRL) passive electrode, were used as reference and ground electrodes, respectively; cf. (www.biosemi/faq/cms&drl.htm). Off-line, trials containing blinks were corrected using a dipole approach (BESA 2000) and EEG activity was re-referenced to average reference. The analysis epoch for the stimulus-locked data started 100 ms prior to cue onset and lasted for a total duration of 1600 ms. For the response-locked analyses, epochs started 1400 ms before the response and lasted for 1600 ms. EEG and EOG activity was filtered (band-pass 0.01-10 Hz), averaged time-locked to cue onset (S-locked data) or to the response onset (R-locked data).

5.1.1.5.1 Behavioral data analysis

As a measure of performance accuracy we chose d' to obtain response bias free estimate of performance by using the z-scores for the hit rate and false alarm rate (for target present trials and for target absent trials, respectively) for each of the two conditions (see also Figure 5.3). For each observer and condition (test and control) we obtained d' estimates and we entered the raw data (trial-by-trial response times) into a spreadsheet containing success trials. For each observer and condition we then calculated a d' estimate for each of the twelve trials. The d' measure was amended to a random number that reflected ceiling performance. The d' measure for each of the twelve

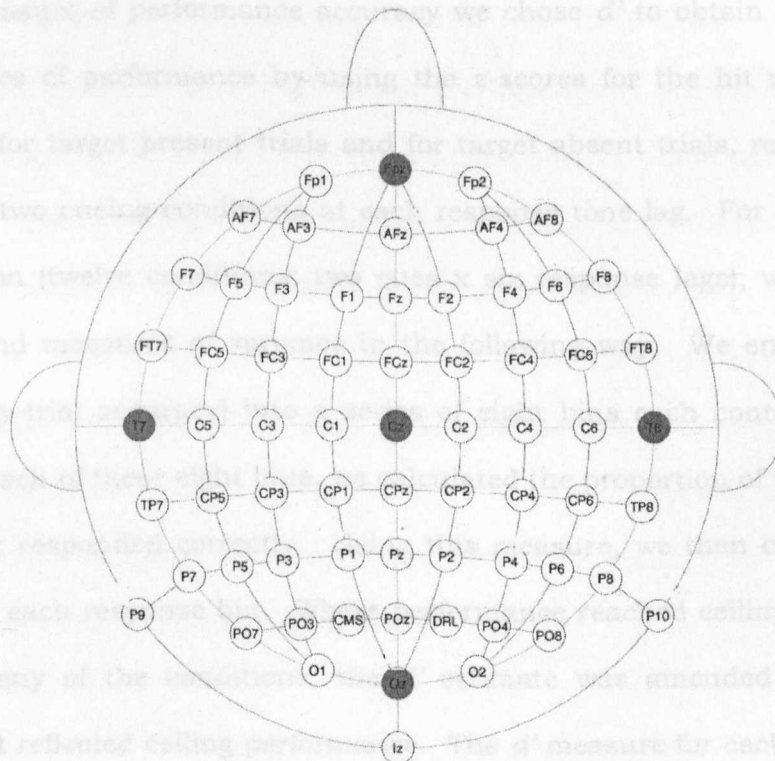


Figure 5.3. Electrode layout for the Biosemi 64 channel system.

In addition, we calculated the LRP for each participant and each experimental condition. To this end, the ERP at recording sites ipsilateral to the response hand was subtracted from the ERP at homologous contralateral recording sites. The resulting difference waveform for each site-pair (e.g., C3/C4) was averaged across hands separately to eliminate any ERP activity unrelated to hand-specific motor activation (Coles, 1989). The term LRP will be exclusively used to describe activity at the C3/C4 site and the term L-ERP to refer to lateralised ERP activity calculated in the same way as the LRP for all other electrode site-pairs.

5.2.1.5 Data Analysis

5.1.1.5.1 Behavioural data analysis

As a measure of performance accuracy we chose d' to obtain response bias free estimates of performance by using the z-scores for the hit rate and false alarm rate (for target present trials and for target absent trials, respectively) for each of the two cueing conditions at each response tone lag. For each observer and condition (twelve conditions: two cues x six response lags), we obtained d' estimates and measures of variance in the following way. We entered the raw data (trial-by-trial accuracy) into a series of eight bins each containing sixteen trials. For each of these eight bins, we calculated the proportion of trials in which the observer responded correctly. Using this measure, we then calculated a d' estimate for each response bin. Where performance reached ceiling levels (100% correct) in any of the conditions, the d' estimate was amended to a random number that reflected ceiling performance. The d' measure for each of the twelve conditions was calculated as the mean d' across these eight bins. For each

condition, the variance was calculated as the variance in d' units across the eight bins. To obtain group data, we calculated the mean d' and variance across all twelve observers for each of the twelve conditions. The d' and variance estimates for each condition were incorporated into the fitting procedures for each observer and for the group mean data.

To measure the effect of attention on processing dynamics the experimental SAT data, for each observer, were fitted with a limited exponential model with three parameters:

$$d'(t) = \lambda * (1 - \exp(-(x - \sigma) / \gamma)) \quad (\text{Equation 5.1})$$

The asymptotic parameter λ reflects the point at which maximal performance accuracy is achieved, γ corresponds to the rate of increase in accuracy to asymptotic performance, measured in the pre-asymptotic section of the function and σ , the intercept of the function marks the point at which accuracy departs from chance performance. As proposed by McElree and Carrasco (1999), we assessed the goodness of fit by evaluating an adjusted- R^2 statistic that represents the amount of variance captured by the model. This procedure was conducted on each individual observer's SAT data to obtain estimates of each of the three parameters, and also on the average data. To test for differences between cueing conditions, we conducted paired-samples t -tests (one-tailed) on the individual asymptote, rate and intercept parameters.

5.2.1.5.2 EEG data analysis

Only trials with a correct response and without EEG or EOG artifacts were included in the EEG data analysis. All signals were averaged separately for each observer and each of the twelve experimental conditions (two cue types and six response tone lags). Mean ERP amplitude was analysed in a 200 ms interval immediately preceding response signal onset using a 200 ms baseline before pre-cue onset. To examine the effects of cueing on the ERP data, statistical analyses were performed by means of Greenhouse-Geisser corrected (Huyhn, 1978) repeated measures analyses of variance (ANOVA) at time intervals (from 80 to 380 ms post-stimulus).

LRP onsets were measured and analysed by applying a jack-knife-based procedure (Miller, Patterson, and Ulrich, 1998). For each condition, grand mean LRP waveforms were calculated by averaging across all participants. Next, grand average LRPs for each of the twelve experimental conditions were computed by omitting from each grand average the ERP data of another participant. LRP onsets were measured in the waveform of each grand average as follows. Because LRP peak amplitudes in S-locked waveforms differed as a function of experimental conditions, a fixed onset threshold was used (Miller et al., 1998). In Experiments 5.1 and 5.3, LRP onsets were measured relative to a 1-100 ms pre-response signal baseline (for Experiment 5.3, this baseline was set to between 600-700 ms for the longer lags) at the point in time when LRP amplitude exceeded 1 μ V. For the shorter lags (30, 60, 120, and 250 ms), the search area for the S-locked LRP was conducted in the 300-800 ms time interval, whereas for the longer lags (500 and 1000 ms), the search area was set to the 800 to 1500 ms time interval. In Experiment 5.2, for the short response lags (30, 60, 120, and

250 ms), LRP onsets were measured relative to a 1-100 ms pre-response signal baseline at the point in time when LRP amplitude exceeded 0.5 μ V, whereas for the longer lags, the signal baseline was set to 600-700 ms. For the shorter lags the search area for the S-locked LRP was conducted in the 300-800 ms time interval, whereas for the longer lags (500 and 1000 ms), the search area was set to the 800 to 1500 ms time interval. Following the recommendations of Miller et al. (1998), effects in the LRP-R interval were obtained using a relative LRP onset criteria (30% of LRP peak amplitude) with waveforms referred to a 200 ms baseline starting 1400 ms before the response. The standard error of the difference between LRP onsets in cued and neutral conditions was determined with the jack-knife method proposed by Miller et al. (1998) and the null hypotheses were tested using repeated-measures ANOVAs with Cue and Lag as factors.

5.2.2 Results

5.2.2.1 Behavioural measures

5.2.2.1.1 Reaction time and Accuracy

Table 5.1 shows mean reaction time (RT) and accuracy (shown as proportion correct) of twelve observers for the cued and neutral conditions at six response lags. A repeated measures ANOVA on the RT data with Cue and Lag as factors showed a significant main effect of Lag ($F(5,55)=821.35$, $p<0.001$), a non-significant main effect of Cue ($F(1,11)=0$, $p>0.05$), and significant Lag x Cue interaction ($F(5,55)=23.27$, $p<0.001$). The same analysis on the accuracy data showed significant main effects of Lag ($F(5,55)=40.75$, $p<0.001$) and Cue

($F(1,11)=14.68$, $p<0.001$), and significant Lag x Cue interaction ($F(5,55)=5.64$, $p<0.001$). Independent analyses of reaction time and accuracy data for Experiment 5.1 therefore indicate that the valid peripheral cue did not modulate reaction times (mean RT 594.71 and 594.64, for cued and neutral conditions respectively), but did enhance performance accuracy (mean proportion correct 91.5 and 88.5, for cued and neutral conditions respectively). The Cue x Lag interaction is due to shorter RT and higher accuracy at short response lags for cued than neutral trials, whereas the cueing effect is reversed in RT, and absent in accuracy, for long response lags (cf. Table 5.1).

Table 5.1. Mean reaction times (ms) and accuracy (proportion correct) for cued and neutral conditions at six response lags (SD in parentheses).

Lag	Cued RT	Neutral RT	Cued Acc (%)	Neutral Acc (%)
30 ms	392 (24)	427 (33)	83 (8.5)	79 (7.8)
60 ms	406 (22)	431 (30)	86 (9.2)	80 (8.6)
120 ms	424 (22)	445 (32)	90 (7.2)	85 (9.3)
250 ms	509 (19)	512 (22)	96 (2.2)	94 (4.8)
500 ms	713 (28)	695 (28)	96 (2.4)	97 (3.0)
1000 ms	1125 (89)	1059 (93)	97 (2.1)	97 (2.3)

5.2.2.1.2 Speed-Accuracy Tradeoff

To examine the dynamics of information processing directly, one must consider conjointly speed and discriminability measures. Figure 5.4 shows the group average data, demonstrating a clear speed-accuracy trade-off effect for the cued and neutral conditions. The adjusted- R^2 was .98 for both the cued and the neutral conditions, indicating good fits to the average group data. The individual data yielded adjusted- R^2 statistics ranging between .43 and .93 for the cued condition and between .59 and .98 for the neutral condition.

Therefore, at maximal processing time, there was no difference in the level of asymptotic accuracy between these two cueing conditions.

Next, to examine the effect of cueing on the temporal dynamics of processing, we turn to the rate and intercept parameters of the SAT function. Figure 5.4 shows no difference in the rate parameter, both 51 ms, for the cued and neutral conditions ($t(11) = -2.82, p < 0.01$), indicating that the valid cue did not modulate the rate at which information was processed. Turning to the intercept parameter, a test on the individual data was significant ($t(11) = -2.82, p < 0.01$), with processing proving to be faster in the cued condition (339 ms) than the neutral condition (371 ms). The SAT data of Experiment 5.1 therefore indicate that the valid peripheral cue did not impair target discriminability, as indicated by a similar asymptote in both cueing conditions. In relation to the rate of cueing, the speed of information processing, no effect was evident in the rate of information processing, however,

an earlier departure from the asymptote was observed in the cued relative to the neutral condition.

Figure 5.4. Average time course data for Experiment 5.1. Accuracy is plotted in d' units as a function of processing time (response lag plus processing time in seconds) for both cued and neutral conditions. The data is fitted with Equation 5.1.

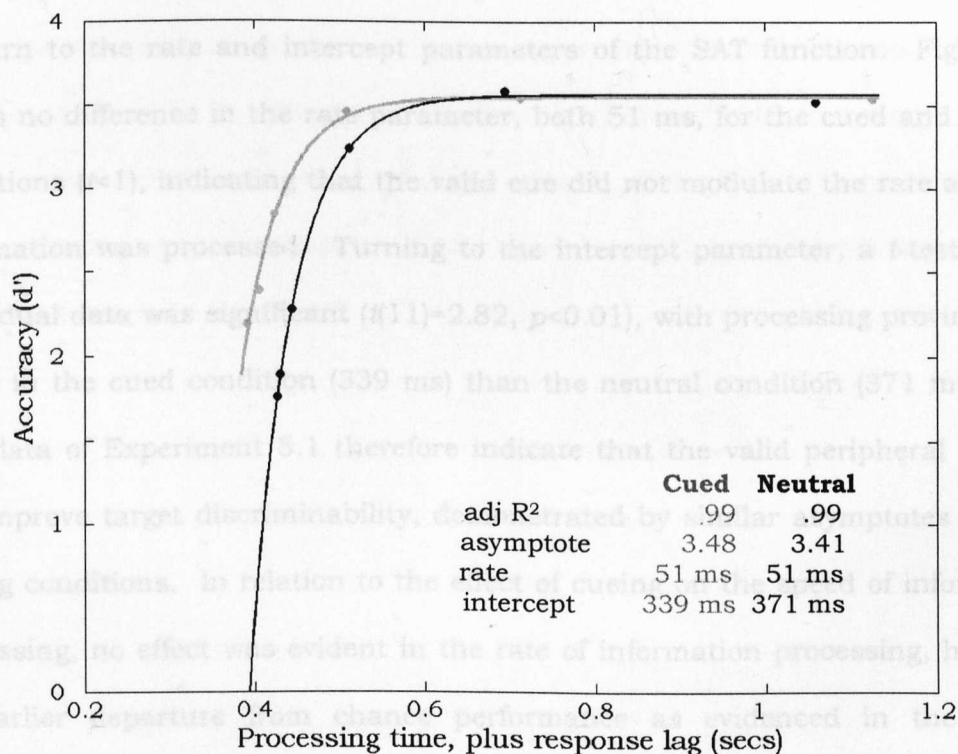


Figure 5.4 shows the values for each of the three parameters used to fit the group data. To examine differences in discriminability between the cued and neutral conditions, we looked at the asymptote, the point at which maximal performance is achieved. The asymptotic parameters of the SAT function showed that accuracy reached asymptotic levels at 3.48 d' units for the cued condition, and 3.41 d' units for the neutral condition, indicating that performance was only marginally higher in the cued condition. A t -test on the individual asymptote parameters was not significant, $t(11)=0.26$, $p>0.05$, one-tailed. Therefore, at maximal processing time, there was no difference in the level of asymptotic accuracy between these two cueing conditions.

Next, to examine the effect of cueing on the temporal dynamics of processing, we turn to the rate and intercept parameters of the SAT function. Figure 5.4 shows no difference in the rate parameter, both 51 ms, for the cued and neutral conditions ($t<1$), indicating that the valid cue did not modulate the rate at which information was processed. Turning to the intercept parameter, a t -test on the individual data was significant ($t(11)=2.82$, $p<0.01$), with processing proving to be faster in the cued condition (339 ms) than the neutral condition (371 ms). The SAT data of Experiment 5.1 therefore indicate that the valid peripheral cue did not improve target discriminability, demonstrated by similar asymptotes in both cueing conditions. In relation to the effect of cueing on the speed of information processing, no effect was evident in the rate of information processing, however, an earlier departure from chance performance as evidenced in the earlier intercept of the SAT function was demonstrated for the cued relative to the neutral condition.

5.2.2.2. Electrophysiological measures

5.2.2.2.1 The Stimulus-locked LRP (S-LRP)

Figure 5.5 depicts, from left to right, the average stimulus-locked LRP for correct responses for cued and neutral conditions at each response tone lag. For each response lag, we used the jack-knife-based scoring method to measure the difference in S-LRP onset latencies between the cued and neutral conditions; Table 5.2 provides the mean onset latencies for cued and neutral conditions at six response lags. A repeated-measures ANOVA on the mean onset latencies, with Cue and Lag as factors, showed a non-significant main effect of Cue ($F(1,11)=0.39$, $p>0.05$), and significant main effect of Lag ($F(5,55)=60.43$, $p<0.001$). The Cue x Lag interaction was not significant ($F(5,55)=1.53$, $p>0.05$). The type of pre-cue, peripheral valid or central neutral therefore had no significant effect on S-LRP latency, indicating that early pre-motoric processing was not modulated by the validity of the pre-cue. The effect of response tone lag was evidenced as an increase in S-LRP mean onset latency as the lag between target and response increased.

Table 5.2. The mean onset-latency (ms) of the stimulus-locked LRP (S-LRP) for cued and neutral trials at six response lags in Experiment 5.1 (*SD* in parentheses).

Lags	Cued	Neutral
30 ms	496 (8.55)	496 (6.40)
60 ms	485 (6.02)	542 (7.51)
120 ms	488 (4.46)	546 (9.39)
250 ms	570 (4.82)	570 (6.38)
500 ms	867 (5.70)	850 (6.34)
1000 ms	1300 (52.63)	1108 (18.87)

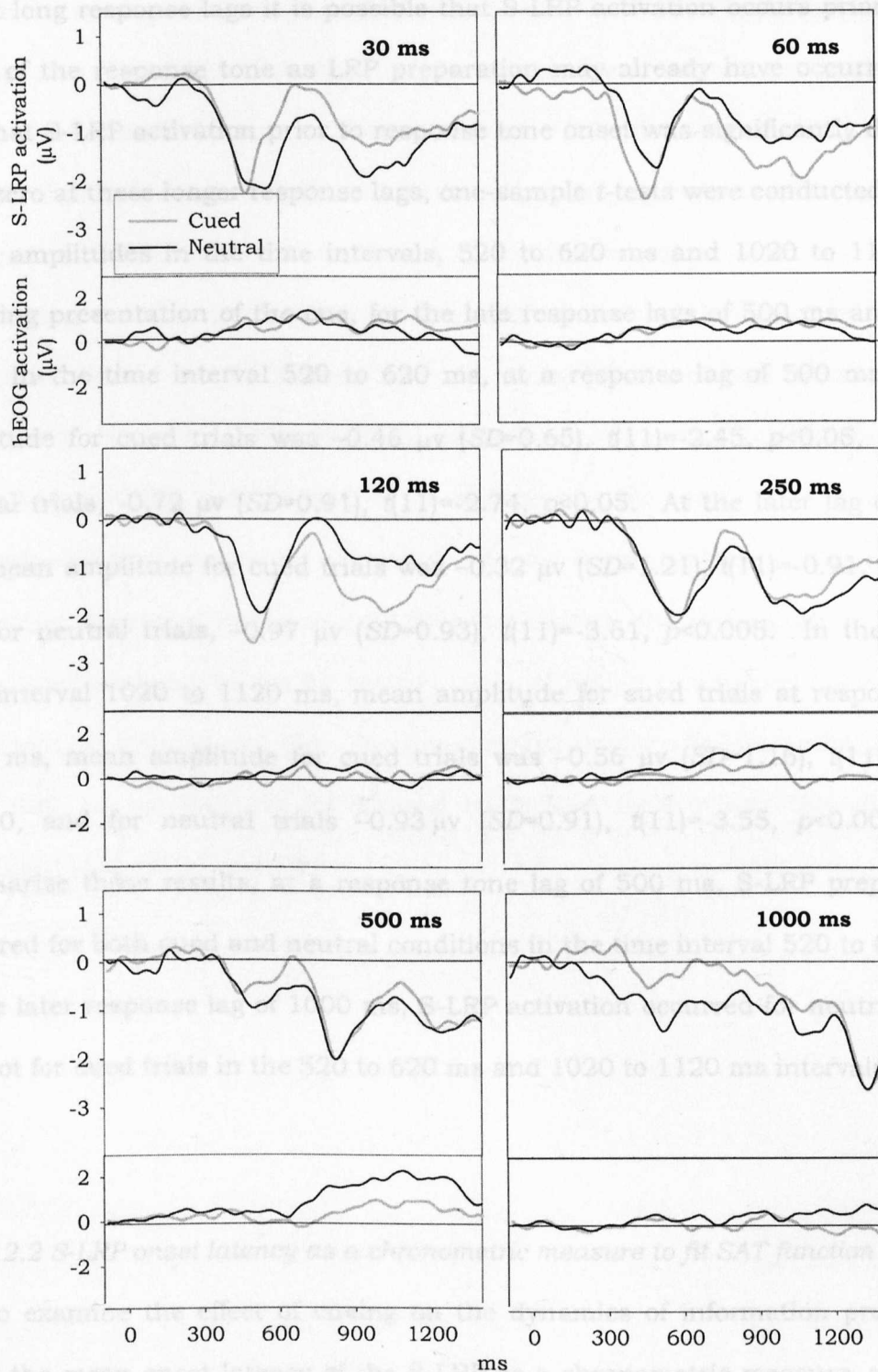


Figure 5.5. The stimulus-locked LRP and HEOG for Experiment 5.1. The top waveform in each panel shows the group average S-LRP for both conditions, cued and neutral, and related HEOG is shown in the bottom waveform, for each response lag, from left-to-right, top to bottom.

At long response lags it is possible that S-LRP activation occurs prior to the onset of the response tone as LRP preparation may already have occurred. To test that S-LRP activation prior to response tone onset was significantly different from zero at these longer response lags, one-sample *t*-tests were conducted on the mean amplitudes in the time intervals, 520 to 620 ms and 1020 to 1120 ms, following presentation of the cue, for the late response lags of 500 ms and 1000 ms. In the time interval 520 to 620 ms, at a response lag of 500 ms, mean amplitude for cued trials was $-0.46 \mu\text{v}$ ($SD=0.65$), $t(11)=-2.45$, $p<0.05$, and for neutral trials, $-0.72 \mu\text{v}$ ($SD=0.91$), $t(11)=-2.74$, $p<0.05$. At the later lag of 1000 ms, mean amplitude for cued trials was $-0.32 \mu\text{v}$ ($SD=1.21$), $t(11)=-0.91$, $p>0.05$, and for neutral trials, $-0.97 \mu\text{v}$ ($SD=0.93$), $t(11)=-3.61$, $p<0.005$. In the S-LRP time interval 1020 to 1120 ms, mean amplitude for cued trials at response lag 1000 ms, mean amplitude for cued trials was $-0.56 \mu\text{v}$ ($SD=1.16$), $t(11)=-1.69$, $p>0.10$, and for neutral trials $-0.93 \mu\text{v}$ ($SD=0.91$), $t(11)=-3.55$, $p<0.005$. To summarise these results, at a response tone lag of 500 ms, S-LRP preparation occurred for both cued and neutral conditions in the time interval 520 to 620 ms. At the later response lag of 1000 ms, S-LRP activation occurred for neutral trials but not for cued trials in the 520 to 620 ms and 1020 to 1120 ms intervals.

5.2.2.2.2 S-LRP onset latency as a chronometric measure to fit SAT function

To examine the effect of cueing on the dynamics of information processing using the mean onset latency of the S-LRP as a chronometric measure, we fitted Equation 5.1 to the data presented in Table 5.2. Since the S-LRP onset latencies were obtained using a jack-knife procedure, we subjected the d' and variance

data to the same procedure (described in Section 5.2.1.5.2). Figure 5.6 shows the average data, fitted with Equation 5.1, for the cued and neutral conditions. The adjusted- R^2 was .76 for the cued condition and .72 for the neutral condition. The individual data fits yielded adjusted- R^2 statistics ranging between .52 and .99 for the cued condition and between .56 and .77 for the neutral condition.

Turning to the parameters of the SAT function, performance reached asymptotic levels at 3.55 and 3.60 d' units for cued and neutral trials respectively. A t -test on the individual asymptotic parameters was not significant ($t(11)=0.46$, $p>0.05$), indicating that the peripheral valid cue did not improve the discriminability of the target stimulus. As for a modulation in the speed of processing, measured here using S-LRP onset latency, the rate parameter was 47 ms for cued trials vs. 81 ms for neutral trials, although this turned out not to be significant ($t(11)=0.20$, $p>0.05$). The intercept parameter was also earlier for cued trials 429 ms vs. 447 ms for neutral trials, but neither was this difference significant ($t(11)=0.04$, $p>0.05$).

Lag ($F(5,55)=5.79, p<0.001$). The Cue \times Lag interaction was also non-significant ($F(5,55)=1.28, p>0.05$). Table 5.3 shows the mean onset latencies of the response-locked LRP for cued and neutral conditions at each of the six response lags. Again, the valid peripheral cue had no significant effect on the onset latency of the LRP-R, indicating that motoric processes were not significantly modulated by the validity of the pre-response cue. The effect of the response tone lag on the LRP-R mean onset latency was likely to lengthen the onset latency as the response lag increased, with the exception of the longer lags.

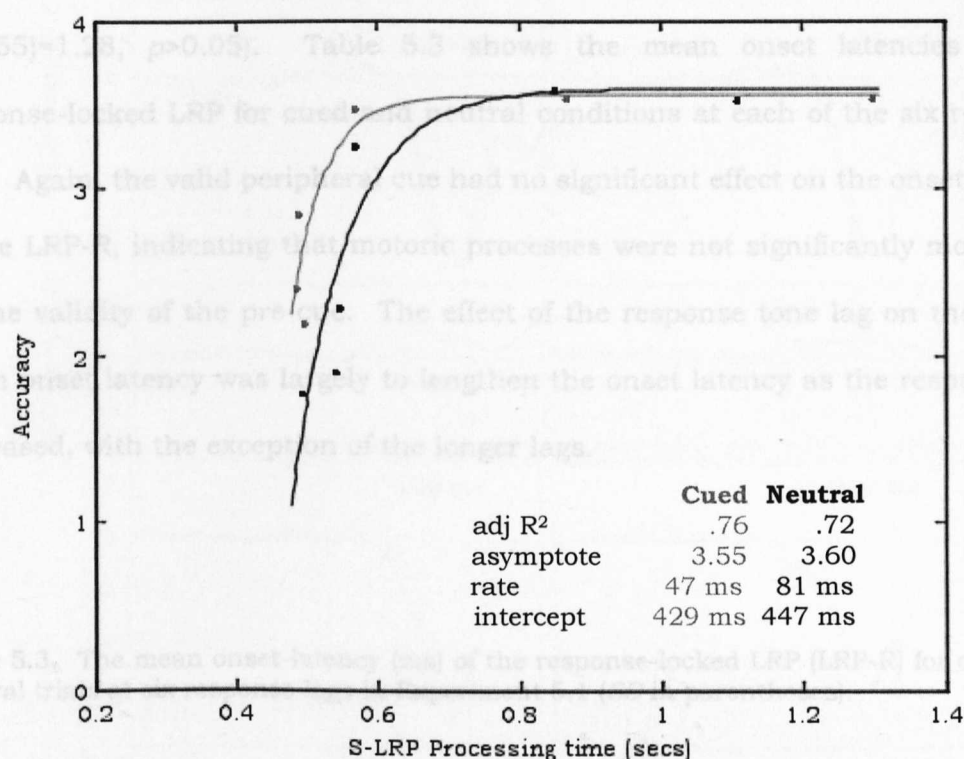


Figure 5.6. Average discrimination accuracy (d' units) as a function of S-LRP latency for Experiment 5.1. Accuracy is plotted in d' units as a function of S-LRP onset latency (response lag plus processing time in seconds) for both cued and neutral conditions and is fitted with Equation 5.1.

60 ms -153 (2.70) -136 (2.31)

120 ms -151 (2.31) -136 (2.51)

250 ms -175 (4.59) -161 (2.61)

500 ms -191 (4.19) -135 (2.57)

5.2.2.3 The Response-locked LRP (LRP-R)

Figure 5.7 depicts, again from left to right, the average response-locked LRP and related HEOG for the correct responses in the cued and neutral conditions, for each response tone lag. Again, we used the jack-knife-based scoring method to measure the difference in LRP-R onset latencies. A repeated-measures ANOVA on the mean onset latencies with Cue and Lag as factors showed a non-significant main effect of Cue ($F(1,11)=0.10, p>0.05$), and significant main effect of

Lag ($F(5,55)=5.79$, $p<0.001$). The Cue x Lag interaction was also non-significant ($F(5,55)=1.28$, $p>0.05$). Table 5.3 shows the mean onset latencies of the response-locked LRP for cued and neutral conditions at each of the six response lags. Again, the valid peripheral cue had no significant effect on the onset latency of the LRP-R, indicating that motoric processes were not significantly modulated by the validity of the pre-cue. The effect of the response tone lag on the LRP-R mean onset latency was largely to lengthen the onset latency as the response lag increased, with the exception of the longer lags.

Table 5.3. The mean onset-latency (ms) of the response-locked LRP (LRP-R) for cued and neutral trials at six response lags in Experiment 5.1 (*SD* in parentheses).

Lag	Cued	Neutral
30 ms	-145 (2.70)	-138 (2.61)
60 ms	-145 (2.70)	-136 (2.31)
120 ms	-151 (2.31)	-136 (2.61)
250 ms	-176 (4.66)	-161 (2.61)
500 ms	-119 (4.18)	-135 (2.67)
1000 ms	-128 (2.31)	-148 (4.42)

5.2.2.4 Event-related potentials (ERPs)

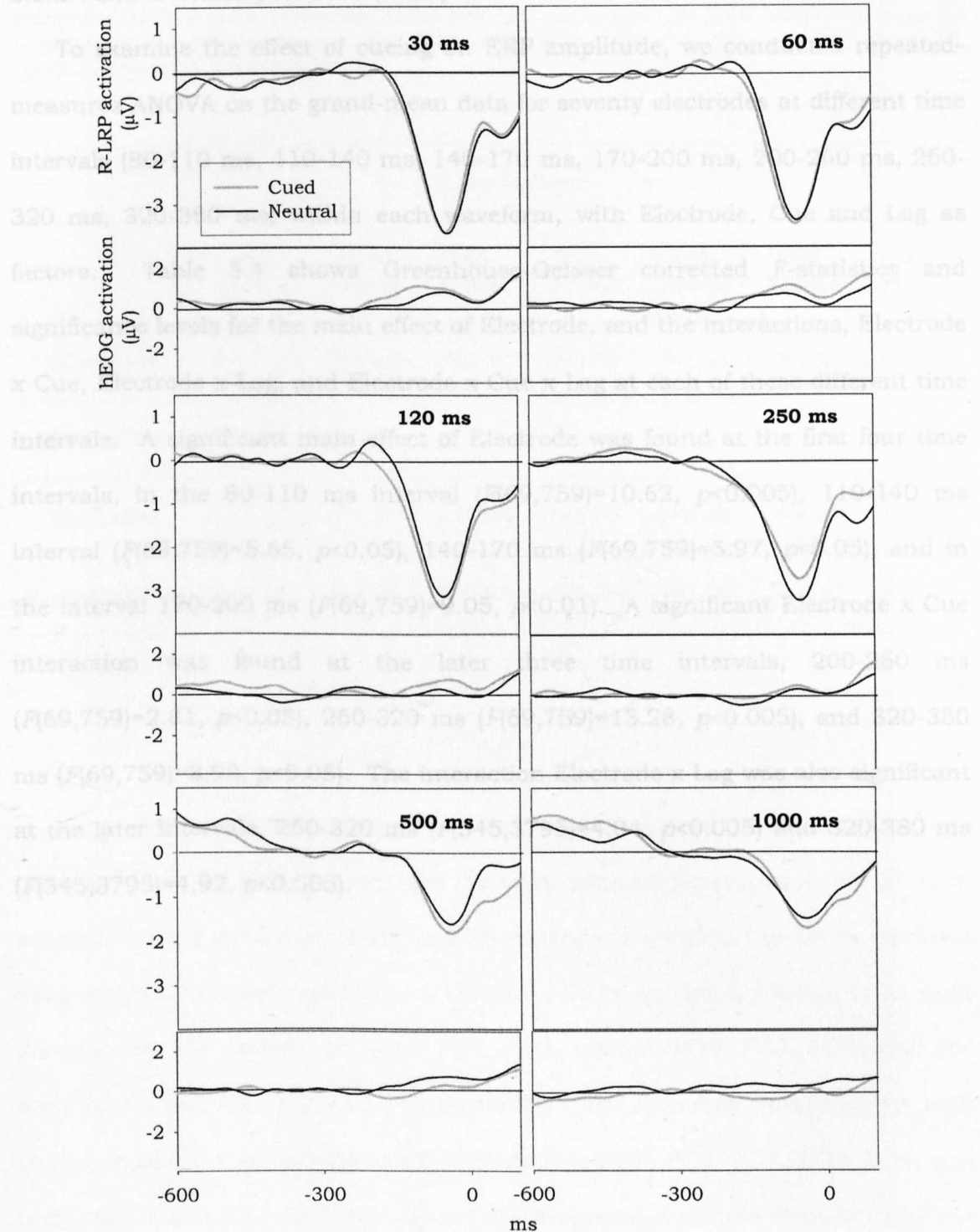


Figure 5.7. The response-locked LRP and HEOG for Experiment 5.1. Each panel, from left-to-right, top-to bottom, shows the cued and neutral LRP-R onset latency, and related HEOG, for each response lag.

5.2.2.4 Event-related potentials (ERPs)

To examine the effect of cueing on ERP amplitude, we conducted repeated-measures ANOVA on the grand-mean data for seventy electrodes at different time intervals (80-110 ms, 110-140 ms, 140-170 ms, 170-200 ms, 200-260 ms, 260-320 ms, 320-380 ms) within each waveform, with Electrode, Cue and Lag as factors. Table 5.4 shows Greenhouse-Geisser corrected F -statistics and significance levels for the main effect of Electrode, and the interactions, Electrode x Cue, Electrode x Lag, and Electrode x Cue x Lag at each of these different time intervals. A significant main effect of Electrode was found at the first four time intervals, in the 80-110 ms interval ($F(69,759)=10.62, p<0.005$), 110-140 ms interval ($F(69,759)=5.65, p<0.05$), 140-170 ms ($F(69,759)=5.97, p<0.05$), and in the interval 170-200 ms ($F(69,759)=9.05, p<0.01$). A significant Electrode x Cue interaction was found at the later three time intervals, 200-260 ms ($F(69,759)=2.81, p<0.05$), 260-320 ms ($F(69,759)=13.28, p<0.005$), and 320-380 ms ($F(69,759)=8.93, p<0.05$). The interaction Electrode x Lag was also significant at the later intervals, 260-320 ms ($F(345,3795)=4.04, p<0.005$) and 320-380 ms ($F(345,3795)=4.92, p<0.005$).

Table 5.4. Greenhouse-Geisser corrected *F*-values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data for Experiment 5.1 over electrode sites PO7, PO8, PO3, PO4, O1, O2, and POZ ($\dagger p > 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.005$).

	Time Interval						
	80-110ms	110-140ms	140-170ms	170-200ms	200-260ms	260-320ms	320-380ms
Electrode	<i>F</i> =10.62 ***	<i>F</i> =5.65 *	<i>F</i> =5.97 *	<i>F</i> =9.05 **	<i>F</i> =1.84 †	<i>F</i> =1.58 †	<i>F</i> =1.33 †
Electx Cue	<i>F</i> =1.26 †	<i>F</i> =0.93 †	<i>F</i> =1.47 †	<i>F</i> =2.46 †	<i>F</i> =2.81 *	<i>F</i> =13.28 ***	<i>F</i> =8.93 ***
Electx Lag	<i>F</i> =1.02 †	<i>F</i> =1.16 †	<i>F</i> =1.06 †	<i>F</i> =0.95 †	<i>F</i> =0.89 †	<i>F</i> =4.04 ***	<i>F</i> =4.92 ***
Eleccx Cue x Lag	<i>F</i> =1.12 †	<i>F</i> =1.04 †	<i>F</i> =0.92 †	<i>F</i> =0.96 †	<i>F</i> =1.06 †	<i>F</i> =1.19 †	<i>F</i> =1.21 †

Since there was no significant Electrode x Cue x Lag interaction, at any time interval within the ERP, we averaged the waveforms across response lags for each separate cueing condition. Further ERP analysis is restricted to seven electrode sites where attention-related effects on visual ERPs are usually found to be most pronounced: the parietal-occipital PO7, PO8, midline POZ, PO3, PO4, and the occipital O1 and O2. Figure 5.8 presents the grand-averaged waveforms for both cueing conditions at parietal-occipital electrodes PO7, PO8, POZ, PO3, PO4, and occipital O1 and O2. The figure shows a differential amplitude modulation of the ERP across these electrodes, as a result of cueing.

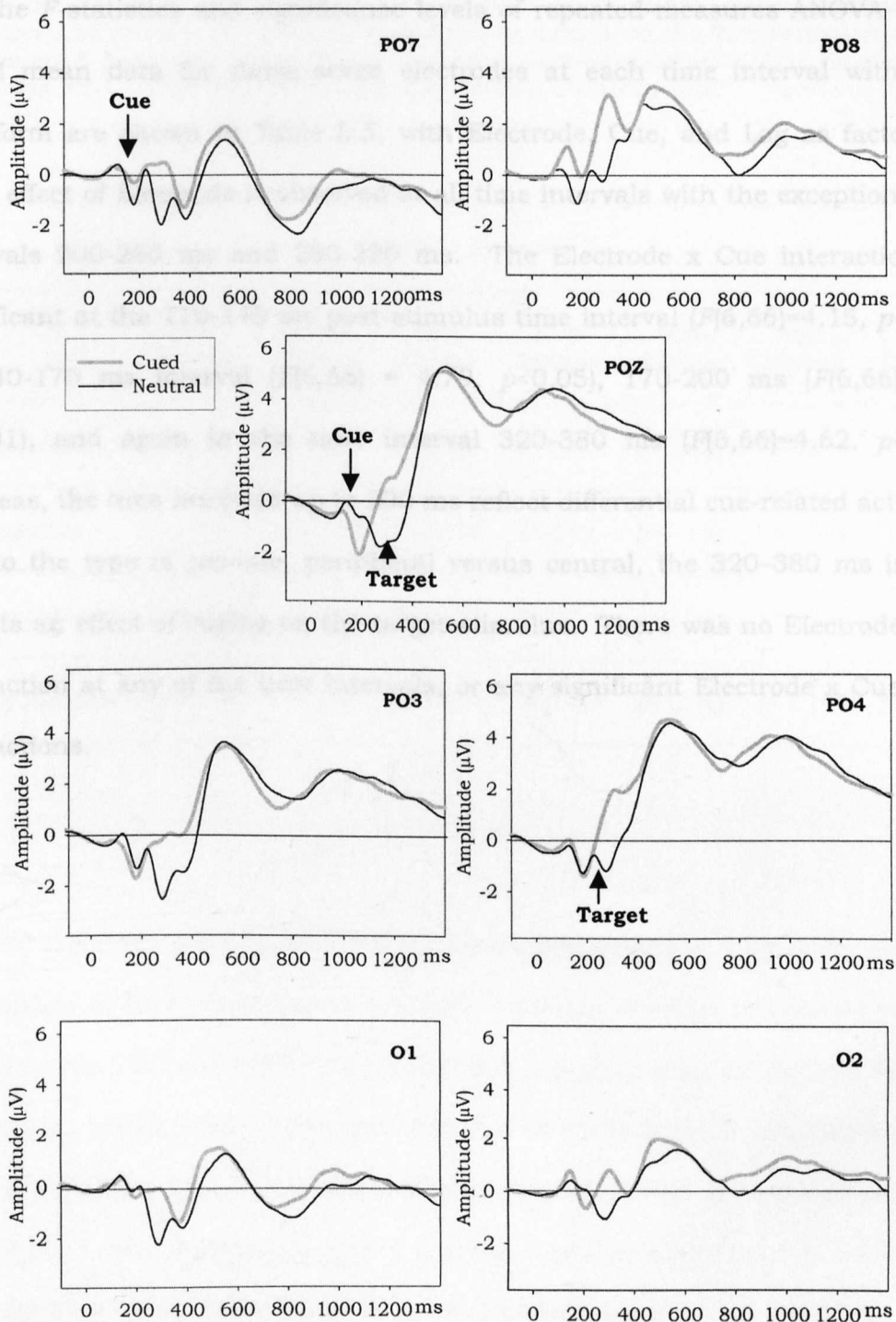


Figure 5.8. Grand-averaged waveforms for cued and neutral conditions of Experiment 5.1. The waveforms depicted are for seven electrode sites involved in visual processing: parietal-occipital electrodes PO7, PO8, midline POZ, PO3, PO4 and occipital electrodes O1 and O2.

The F -statistics and significance levels of repeated-measures ANOVA on the grand mean data for these seven electrodes at each time interval within the waveform are shown in Table 5.5, with Electrode, Cue, and Lag as factors. A main effect of Electrode is observed at all time intervals with the exception of the intervals 200-260 ms and 260-320 ms. The Electrode x Cue interaction was significant at the 110-140 ms post-stimulus time interval ($F(6,66)=4.15, p<0.05$), at 140-170 ms interval ($F(6,66) = 4.79, p<0.05$), 170-200 ms ($F(6,66)=5.91, p<0.01$), and again in the time interval 320-380 ms ($F(6,66)=4.62, p<0.05$). Whereas, the time intervals up to 200 ms reflect differential cue-related activation due to the type of pre-cue, peripheral versus central, the 320–380 ms interval reflects an effect of cueing on the target stimulus. There was no Electrode x Lag interaction at any of the time intervals, or any significant Electrode x Cue x Lag interactions.

Table 5.5. Greenhouse-Geisser corrected *F*-values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on the ERP amplitude data for Experiment 5.1 on seven electrode sites: occipital electrodes O1 and O2; parietal-occipital electrodes PO3, PO4, PO7 PO8, and the midline electrode POZ. (†*p*>0.05, **p*<0.05, ***p*<0.01).

	Time Interval						
	80-110ms	110-140ms	140-170ms	170-200ms	200-260ms	260-320ms	320-380ms
Electrode	<i>F</i> =6.07**	<i>F</i> =3.69*	<i>F</i> =4.48*	<i>F</i> =4.35*	<i>F</i> =2.66†	<i>F</i> =3.44†	<i>F</i> =4.70*
Electx Cue	<i>F</i> =0.95†	<i>F</i> =4.15*	<i>F</i> =4.79*	<i>F</i> =5.91**	<i>F</i> =1.12†	<i>F</i> =1.46†	<i>F</i> =4.62*
Electx Lag	<i>F</i> =0.96†	<i>F</i> =1.52†	<i>F</i> =1.33†	<i>F</i> =0.71†	<i>F</i> =0.74†	<i>F</i> =1.62†	<i>F</i> =1.67†
ElectxCuexLag	<i>F</i> =1.23†	<i>F</i> =1.85†	<i>F</i> =1.69†	<i>F</i> =1.48†	<i>F</i> =1.54†	<i>F</i> =1.18†	<i>F</i> =1.17†

The Electrode x Cue interaction is depicted in Figure 5.8 for each electrode; an increase in ERP amplitude in the cued condition is evident at lateral parietal-occipital sites PO3 and PO4 on the N190 component and at the midline electrode POZ on the N200. The reverse pattern, that is an increase in amplitude for the neutral condition relative to the cued condition, is evident at electrode sites PO7 and PO8. This differential ERP activation reflects a differential stimulation between the different cue types, peripheral presentation for the valid cue versus central presentation for the neutral cue. Figure 5.9 illustrates the difference in mean amplitudes between these cue types, peripheral valid versus central neutral, at each of these electrode sites in the 170-200 ms time interval

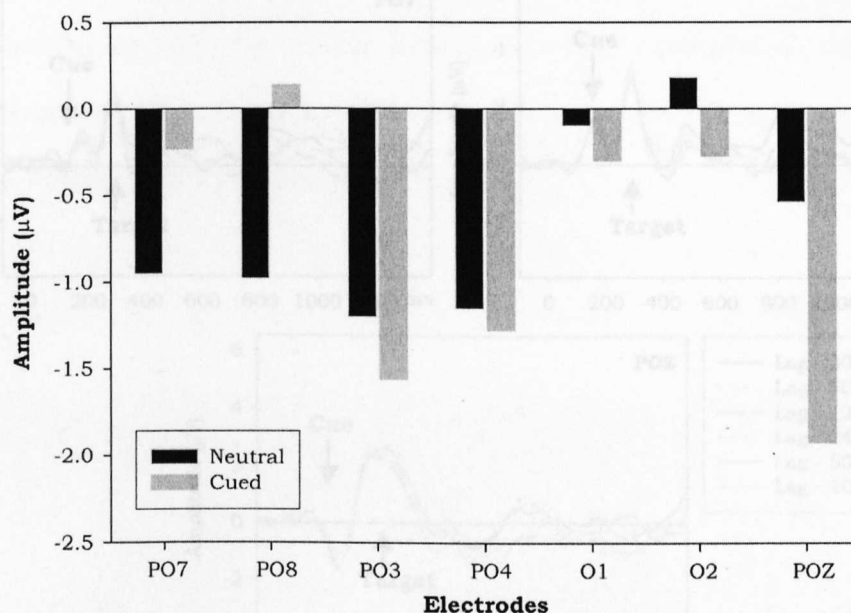


Figure 5.9. Mean amplitude (measured in μV) at electrodes PO7, PO8, PO3, PO4, O1, O2, and POZ in the time interval 170-200 ms post-cue for the cued and neutral conditions of Experiment 5.1.

To examine more closely the time course of the differential activation resulting between different cue types, we calculated difference waves (Cued(ERP activity) minus Neutral(ERP activity)) for each response lag. Figure 5.10 presents these difference waveforms at each electrode site, and clearly depicts cue-related activity particularly around 170 ms, maximal over lateral occipito-temporal recording sites, illustrating the difference in activation between the peripheral cue and central cue. Target-related activation is prominent between 320-380 ms, and is expressed as more positive activity for the valid cue relative to the neutral cue, maximal at electrode Pz. This differential target-related activity between the valid and neutral cue, however, might be due to neuronal refractoriness. However, since the size and location of the pre-cue and target are non-overlapping, this explanation seems unlikely.

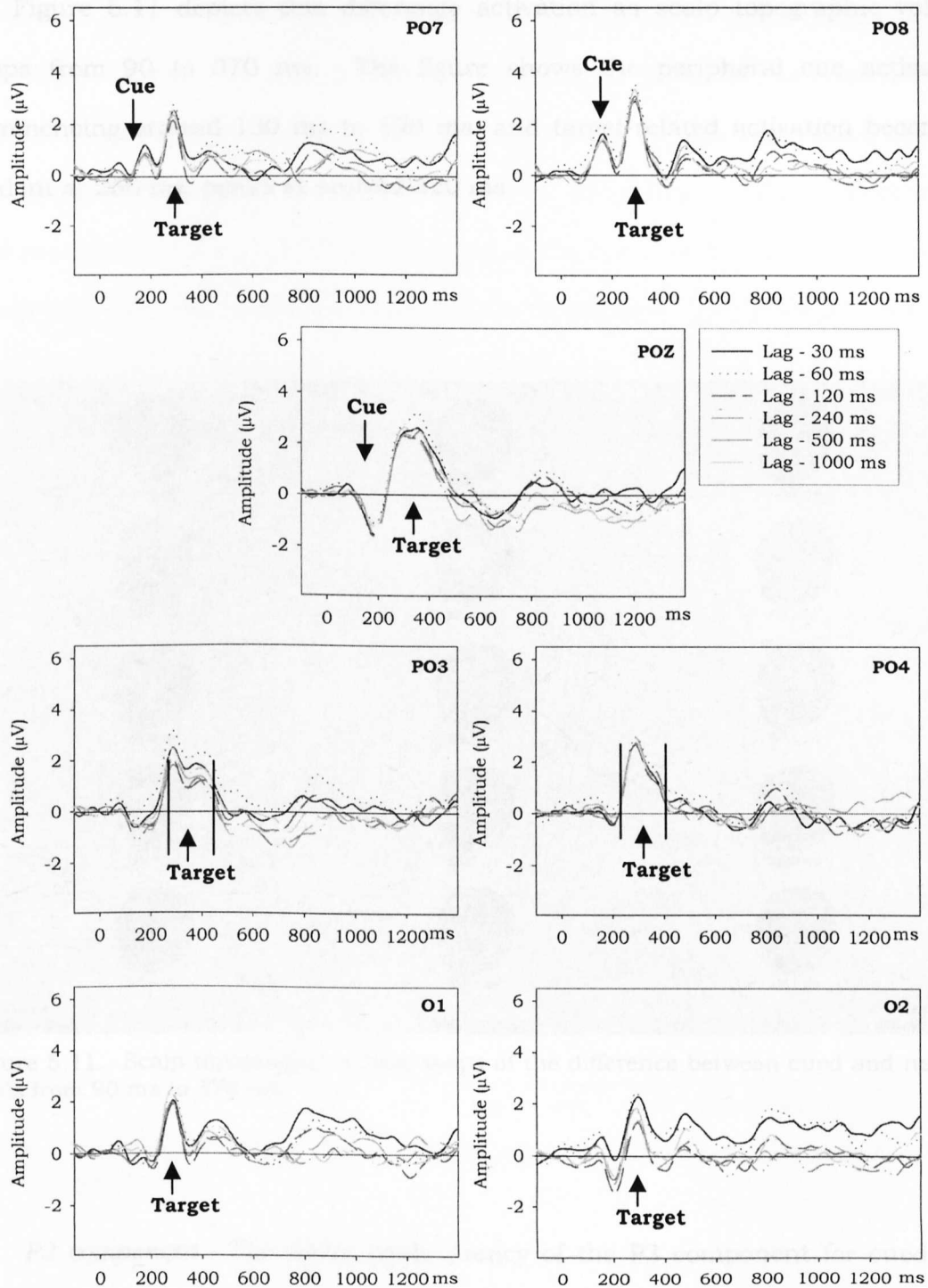


Figure 5.10. Difference waveforms for Cued - Neutral conditions of Experiment 5.1 at each response lag for seven electrode sites (PO7, PO8, POZ, PO3, PO4, O1, and O2). These difference waves were obtained by subtracting the ERP for the neutral condition from the ERP for the cued condition for each response lag.

Figure 5.11 depicts this difference activation as scalp topographic voltage maps from 90 to 370 ms. The figure shows the peripheral cue activation commencing around 130 ms to 190 ms, and target-related activation becoming evident at 250 ms, peaks at around 320 ms.

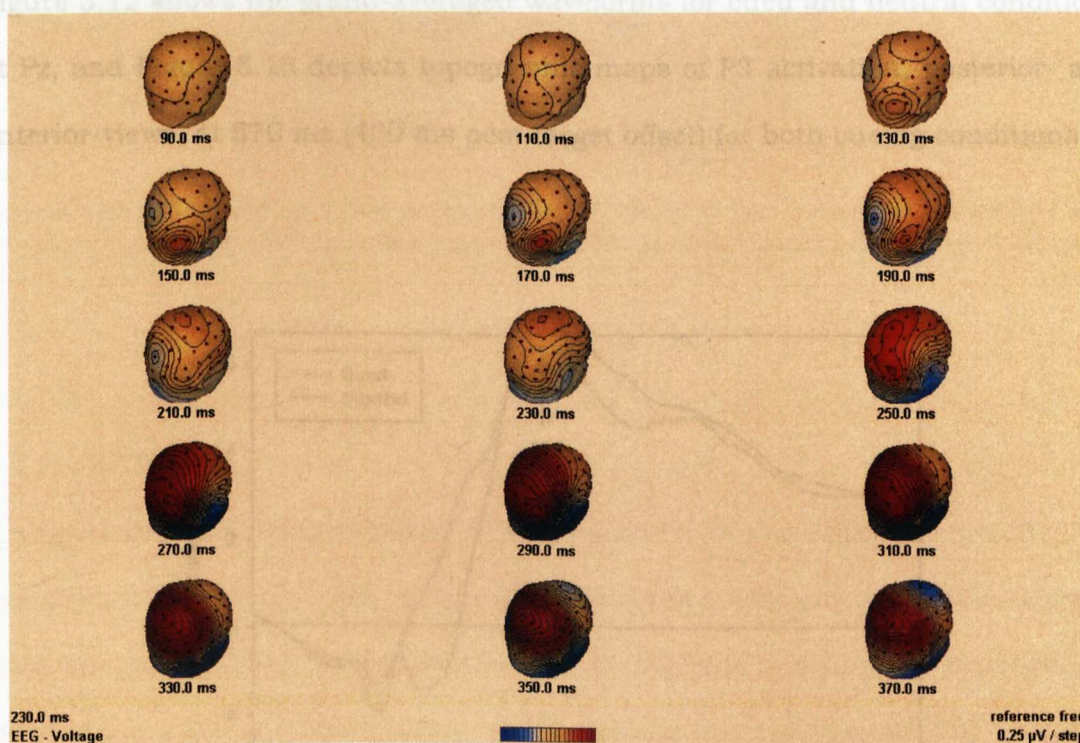


Figure 5.11. Scalp topographic voltage maps of the difference between cued and neutral trials from 90 ms to 370 ms.

P3 component. The mean peak latency of the P3 component for cued and neutral trials was 536.68 ($SD=75.88$) and 565.08 ms ($SD=82.42$), respectively. To analyse the peak latency of this component for cued and neutral trials at electrode Pz, we conducted a repeated measures ANOVA with Lag and Cue as

factors. Results of the ANOVA showed a significant main effect of Lag ($F(5,55)=4.90, p<0.01$), demonstrating a longer peak P3 latency as response tone lag increases (with the exception of the longer lags). A significant main effect of Cue ($F(1,11)=6.64, p<0.05$) was also observed, indicating a reliable effect of the cue resulting in an earlier P3 peak latency in the cued condition than the neutral condition, (but no significant Lag x Cue interaction ($F(5,55)=1.08, p<0.05$)). Figure 5.12 shows the grand-averaged waveforms for cued and neutral conditions at Pz, and Figure 5.13 depicts topographic maps of P3 activation, posterior- and anterior-views, at 570 ms (400 ms post target offset) for both cueing conditions.

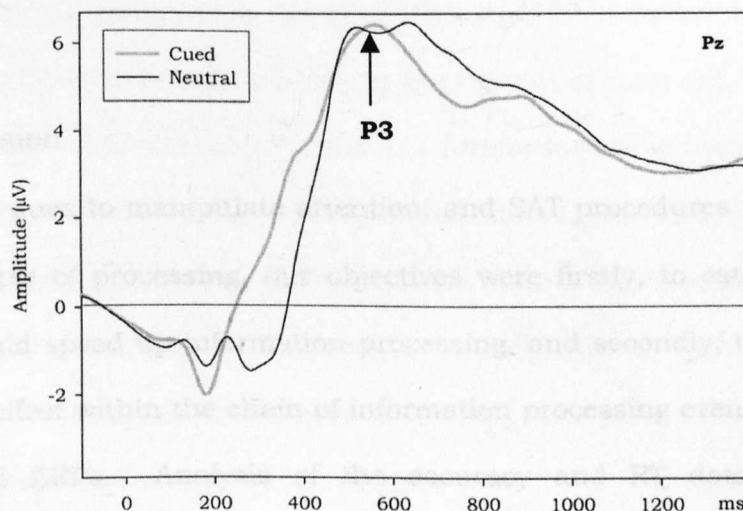


Figure 5.12. Grand-averaged waveform depicting the P3 component for cued and neutral trials at electrode site Pz.

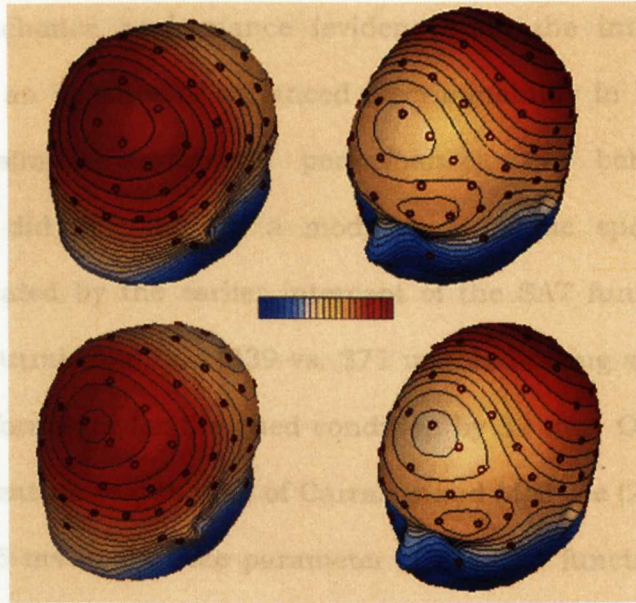


Figure 5.13. Scalp topographic voltage maps of P3 activation for cued (top) and neutral (bottom) trials. Activation is in the time range 570 ms and reference amplitude 0.80 μ V.

5.2.3 Discussion

Using pre-cues to manipulate attention, and SAT procedures to examine the full time course of processing, our objectives were firstly, to establish whether attention would speed up information processing, and secondly, to establish the locus of any effect within the chain of information processing events by analysing the LRP and ERPs. Analysis of the accuracy and RT data suggest that participants followed the SAT procedure, as instructed, in that both accuracy and RT increased as a function of an increase in the response tone lag. This is also supported by the significant lag effect on the onset latency of the LRP.

An attentional modulation in the speed of information processing can be realised as either acceleration in the rate at which information is processed (evidenced on the rate parameter of the SAT function) and/or as an earlier

departure from chance performance (evidenced on the intercept parameter). While there was an absence of enhanced discriminability in the cued condition (evidenced by similar asymptotic performance), the behavioural data of Experiment 5.1 did demonstrate a modulation in the speed of information processing, indicated by the earlier intercept of the SAT function for the cued, relative to the neutral condition (339 vs. 371 ms), indicating an earlier departure from chance performance for the cued condition by 32 ms. Our result contrasts with that of the feature search task of Carrasco and McElree (2001) who observed a difference of 45 ms in the rate parameter of the SAT function. However, this difference reflects the average difference across three conditions used by Carrasco and McElree, incorporating conditions with displays of different set sizes, therefore a direct comparison between the results reported here and their corresponding findings is unavailable. In the current experiment, we did not find a difference in the rate parameter of the SAT function, finding instead a difference in the intercept. However, Carrasco and McElree suggest that the difference in processing time found between their cued and neutral conditions as a difference in rate could also have been expressed as a difference in the intercept of the function, with little loss in the quality of the fits. Our conclusion therefore that cueing the location of an upcoming target modulated the temporal dynamics of processing is consistent with that of Carrasco et al.

What then is the locus of this attentional effect in the information-processing stream? As outlined earlier, analysis of the SAT function does not provide information regarding the particular processes that are speeded up by attention, so in an effort to establish the locus of this effect, I turn from the behavioural data to the EEG data. The effect of cueing on the ERP waveforms at parietal-

occipital and occipital sites firstly became evident between 110-200 ms, and again 320-380 ms. In the earlier intervals, between 110-200 ms, this difference in activation reflects cue-related activity, rather than target-related activity. We did not find an effect of cueing on any of the intervals in which early target-related P1 or N1 would be observed. The later effect of cueing in the 320-380 ms interval, likely to reflect target-related activity, occurs around 200 ms following target onset. However, the types of cues employed in this experiment, peripheral and central, resulted in differential stimulation, and thus incur difficulties in interpreting the resulting differential effects on the ERP waveforms due to the overlapping cue-related and target-related activity in sensory ERP components. As a result, no latency analysis of ERP components was possible.

In relation to the S-LRP component of the LRP, that is the pre-motoric component, we found no difference in onset latency between cueing conditions. Neither did we observe a difference in the onset latency of the LRP-R, the motoric component of the LRP, indicating that motoric processes were not modulated by the informativeness of the peripheral valid cue, relative to the neutral cue. It is therefore possible that the LRP measure was not sufficiently sensitive to tease apart the benefits derived from knowing the target location. The failure to find any modulation on the onset latency of the S-LRP or LRP-R also incurs difficulty in reconciling this data to the SAT data (although analysing the reaction time data independent of accuracy also failed to find a difference between cueing conditions). Yet, we used the S-LRP onset latency data as an alternative chronometric measure in which to fit the SAT function, however statistical testing showed no significant differences between cue types in any of the three SAT parameters. As a chronometric measure, the sensitivity of the LRP to noise has

been acknowledged. However, a difference between cueing conditions was instead found on the peak latency of the P3 component, with an earlier latency for the cued condition relative to the neutral condition. A modulation on the P3 component suggests that the locus of the cueing effect is at the level of stimulus categorisation, reflected by the earlier peak latency in the cued condition (Verleger, 1997). In other words, attention accelerated categorisation-related processes.

In establishing the locus of the effect of attention on the speed of information processing, the current experiment provides limited information due to the difficulties mentioned. In Experiment 5.2, I adapt the cueing procedure to permit comparison between valid and invalid peripheral pre-cues on the ERP waveforms in the hope of establishing the locus of the temporal benefit delivered by the valid pre-cue.

5.3 Experiment 5.2

In Experiment 5.1, we observed a speeding up of visual information processing with attention, expressed on the SAT function as an earlier departure from chance performance, expressed as an earlier intercept in the cued condition relative to the neutral condition. Due to the differential stimulation between the cued and neutral conditions, in that the valid cue was presented peripherally and adjacent to the target, and the neutral cue was presented centrally, we were limited in interpreting the effects of attention on the ERP data of Experiment 5.1. Here, in Experiment 5.2 the cueing procedure is modified to observe the effects of attention on the SAT function and ERPs via the presentation of a valid peripheral

cue and an invalid peripheral cue. Additionally, we modified the spatial arrangement in which cues and targets could appear; here, a cue and target were always presented along the same axis, vertical or horizontal. In simplifying the display in this way, we remove the differential stimulation observed in Experiment 5.1 using the central neutral cue, and observe any potential differences between the vertical and horizontal meridians. Using an invalid cue, as opposed to a neutral cue, may also lead to a greater decrease in performance, further enhancing any attentional effects.

5.3.1 Method

5.3.1.1 Participants

Participants were twelve students from the University of Glasgow with normal, or corrected to normal vision, paid for their participation (mean age of 22, 5 males, 12 right-handed).

5.3.1.2 Stimuli and Procedure

The stimuli and procedure were identical to those used in Experiment 5.1 except for a modification of the cueing procedure and the display. The central neutral and peripheral valid cues from Experiment 5.1 were replaced by peripheral cues that accurately indicated the location of the upcoming target on half of the trials (valid cue) and away from the target on the remaining half (invalid cue). The spatial arrangement was also modified so that on any given trial the target location was restricted to one of two positions along the vertical or

the horizontal axis. As before, the observer's task was to indicate the orientation of the target gabor patch according to the speed-accuracy trade-off procedure.

5.3.2 Results

5.3.2.1 Behavioural data

5.3.2.1.1 Reaction time and accuracy

Table 5.6 presents mean RTs and performance accuracy for valid and invalid conditions at each of the six response lags. A repeated measures ANOVA on the reaction time data with Cue and Lag as factors showed a significant main effect of Lag ($F(5,55)=479.29$, $p<0.001$), a non-significant main effect of Cue ($F(1,11)=1.46$, $p>0.05$), and non-significant Lag x Cue interaction ($F(5,55)= 2.02$, $p>0.05$). The same analysis on the accuracy data showed significant main effects of Lag ($F(5,55)=45.78$, $p<0.001$) and Cue ($F(1,11)=42.45$, $p<0.001$), and a significant Lag x Cue interaction ($F(5,55)=11.50$, $p<0.001$). Independent analyses of the reaction time and accuracy data indicate that the validity of the cue did not significantly modulate reaction time (M (valid vs. invalid) = 574 ms and 579 ms) but did improve performance accuracy (M (valid vs. invalid) = 90.5% and 85.8% correct). The significant Cue x Lag interaction is due to shorter RT and higher accuracy at short response lags for valid than invalid trials, whereas the cueing effect is reversed in RT, and absent in accuracy, for long response lags (cf. Table 5.6).

Table 5.6. Mean reaction times and performance accuracy for valid and invalid conditions at six response lags.

Lag	Valid RT	Invalid RT	Valid Acc (%)	Invalid Acc (%)
30 ms	382 (27)	396 (34)	81 (7.2)	73 (8.4)
60 ms	398 (22)	404 (32)	86 (7.4)	76 (7.5)
120 ms	411 (28)	422 (34)	88 (6.9)	83 (9.0)
250 ms	491 (20)	501 (22)	95 (2.7)	93 (4.8)
500 ms	695 (38)	688 (36)	96 (3.8)	95 (4.7)
1000 ms	1071 (97)	1063 (104)	96 (2.7)	95 (3.1)

5.3.2.1.2 Speed-Accuracy Trade-off

The individual and group average data for Experiment 5.2 were again fitted with Equation 5.1. Fits of the group data yielded adjusted- R^2 values of .99 for both the valid and invalid conditions. Fitting Equation 5.1 to the SAT functions of individual observers yielded adjusted- R^2 values that ranged from between .63 to .97 for the valid condition and .56 to .99 for the invalid conditions. Figure 5.14 depicts the outcome of this fitting procedure on the group data for the valid and invalid conditions. Looking at the asymptotic parameters for the two cue types suggests no effect of attention on overall discriminability at maximal processing time; asymptotic performance was achieved at 3.55 d' units in the valid condition and a d' of 3.57 in the invalid condition, $t(11)=0.77$, $p>0.05$, one-tailed).

5.3.2.2 Electrophysiological Measures

5.3.2.2.1 The S-LRP

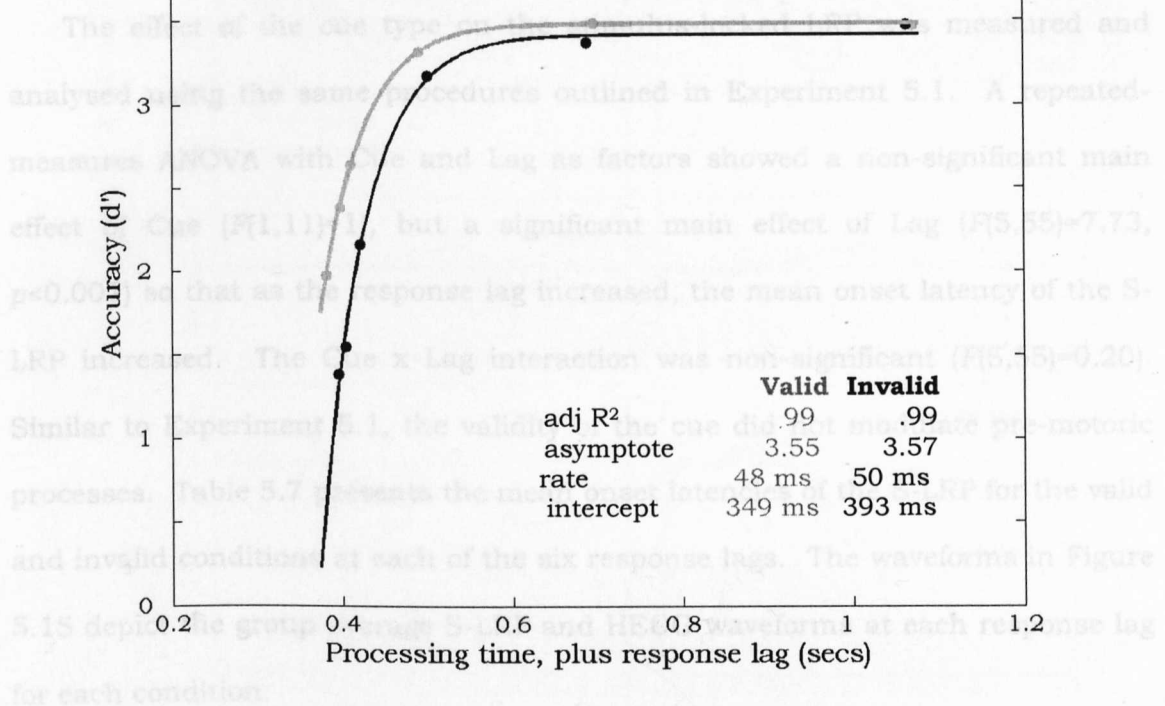


Figure 5.14. Average time course data for Experiment 5.2. Accuracy is plotted in d' units as a function of processing time (response lag plus processing time in seconds) for both valid and invalid conditions. The data is fitted with Equation 5.1.

Table 5.7. The mean onset latency (M) of the stimulus-locked LRP (S-LRP) for the valid and invalid conditions at six response lags in Experiment 5.2 (SD shown in parentheses).

As in Experiment 5.1, to assess the effect of cueing on the dynamics of processing we examine the rate and intercept parameters. No significant difference was found in the rate parameters between cue types, with the valid condition only slightly faster at 48 ms than the invalid condition at 50 ms ($t(11)=0.18$, $p>0.05$). An analysis of the intercept parameters revealed that performance increased from chance levels earlier in the valid condition (349 ms) than the invalid condition (393 ms), a difference of 44 ms ($t(11)=2.36$, $p<0.01$), indicating that the validity of the pre-cue enhanced processing speed.

5.3.2.2 Electrophysiological Measures

5.3.2.2.1 The S-LRP

The effect of the cue type on the stimulus-locked LRP was measured and analysed using the same procedures outlined in Experiment 5.1. A repeated-measures ANOVA with Cue and Lag as factors showed a non-significant main effect of Cue ($F(1,11)=1$), but a significant main effect of Lag ($F(5,55)=7.73$, $p<0.001$) so that as the response lag increased, the mean onset latency of the S-LRP increased. The Cue x Lag interaction was non-significant ($F(5,55)=0.20$). Similar to Experiment 5.1, the validity of the cue did not modulate pre-motoric processes. Table 5.7 presents the mean onset latencies of the S-LRP for the valid and invalid conditions at each of the six response lags. The waveforms in Figure 5.15 depict the group average S-LRP and HEOG waveforms at each response lag for each condition.

Table 5.7. The mean onset-latency (ms) of the stimulus-locked LRP (S-LRP) for the valid and invalid conditions at six response lags in Experiment 5.2 (*SD* shown in parentheses).

Lag	Valid	Invalid
30 ms	450 (8.52)	409 (20.84)
60 ms	470 (7.13)	376 (13.38)
120 ms	449 (10.16)	535 (16.37)
250 ms	532 (7.80)	579 (6.48)
500 ms	852 (4.86)	862 (6.18)
1000 ms	1038 (13.39)	1158 (137.55)

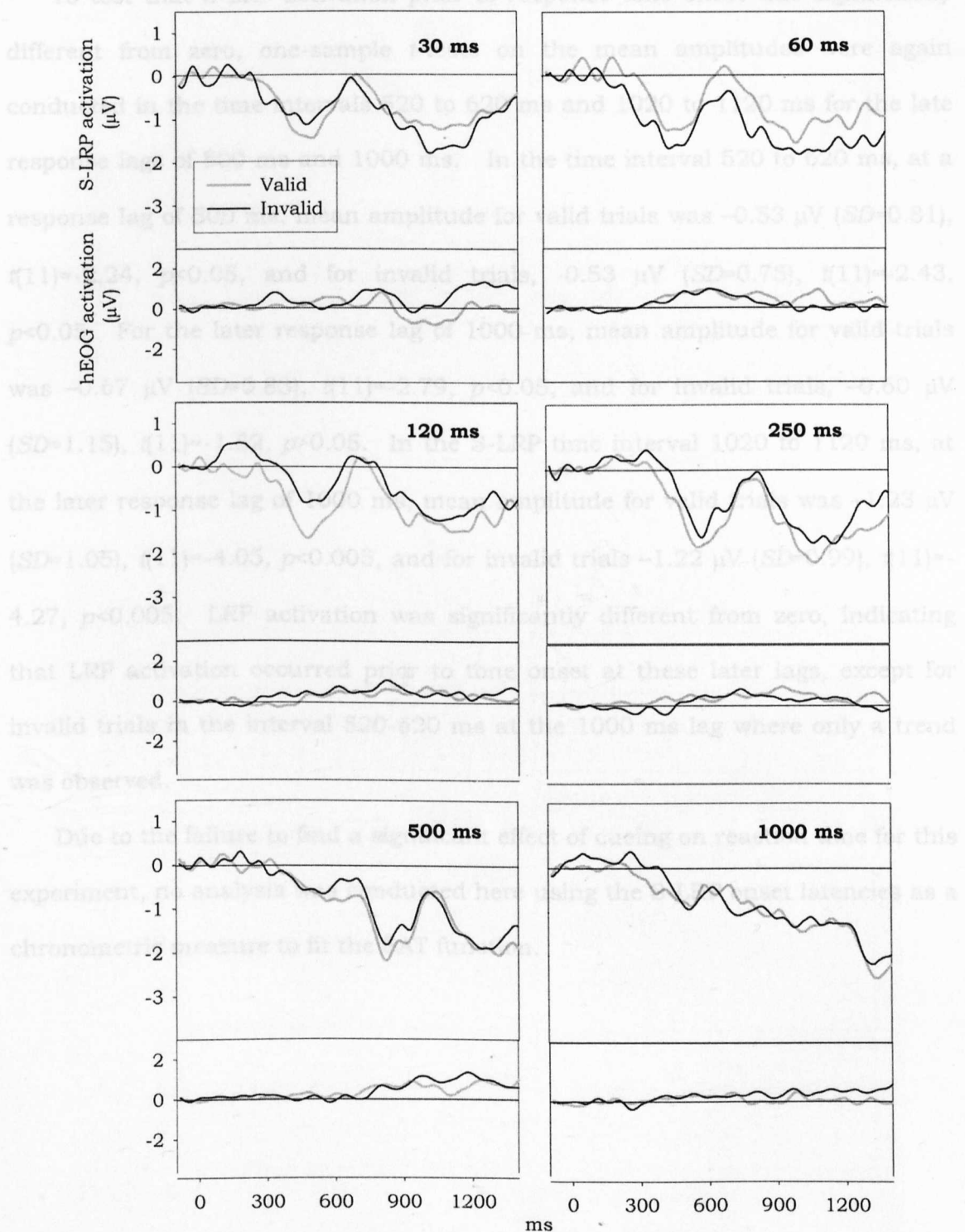


Figure 5.15. The stimulus-locked LRP and HEOG for Experiment 5.2. Each panel, from left-to-right, top-to bottom, shows the valid and invalid S-LRP onset latency, and HEOG, for each response lag.

To test that S-LRP activation prior to response tone onset was significantly different from zero, one-sample *t*-tests on the mean amplitudes were again conducted in the time intervals 520 to 620 ms and 1020 to 1120 ms for the late response lags of 500 ms and 1000 ms. In the time interval 520 to 620 ms, at a response lag of 500 ms, mean amplitude for valid trials was $-0.53 \mu\text{V}$ ($SD=0.81$), $t(11)=-2.24$, $p<0.05$, and for invalid trials, $-0.53 \mu\text{V}$ ($SD=0.75$), $t(11)=-2.43$, $p<0.05$. For the later response lag of 1000 ms, mean amplitude for valid trials was $-0.67 \mu\text{V}$ ($SD=0.83$), $t(11)=-2.79$, $p<0.05$, and for invalid trials, $-0.60 \mu\text{V}$ ($SD=1.15$), $t(11)=-1.82$, $p>0.05$. In the S-LRP time interval 1020 to 1120 ms, at the later response lag of 1000 ms, mean amplitude for valid trials was $-1.23 \mu\text{V}$ ($SD=1.05$), $t(11)=-4.05$, $p<0.005$, and for invalid trials $-1.22 \mu\text{V}$ ($SD=0.99$), $t(11)=-4.27$, $p<0.005$. LRP activation was significantly different from zero, indicating that LRP activation occurred prior to tone onset at these later lags, except for invalid trials in the interval 520-620 ms at the 1000 ms lag where only a trend was observed.

Due to the failure to find a significant effect of cueing on reaction time for this experiment, no analysis was conducted here using the S-LRP onset latencies as a chronometric measure to fit the SAT function.

5.3.2.2.2 The LRP-R

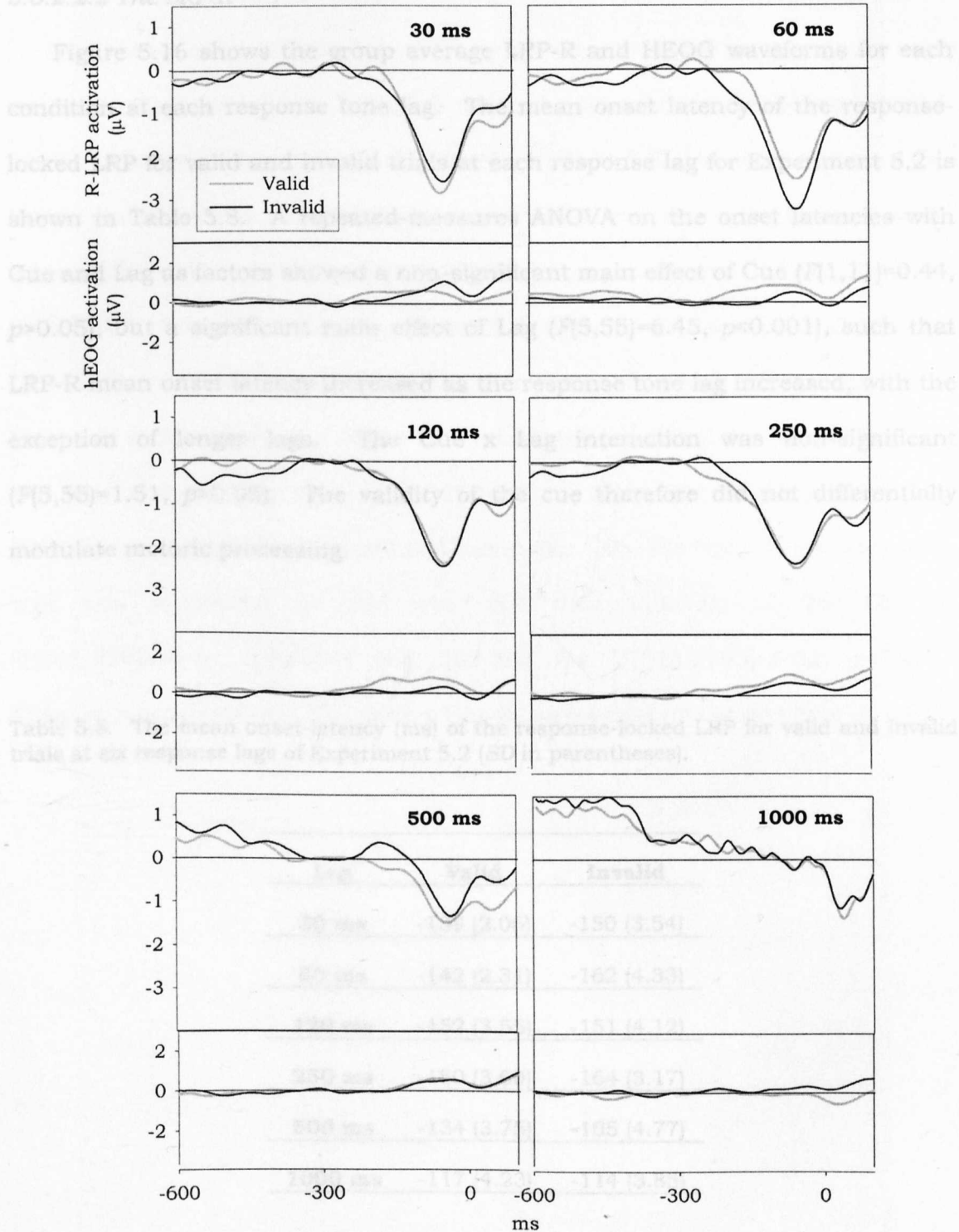


Figure 5.16. The response-locked LRP and HEOG for Experiment 5.2. From left-to-right, top-to-bottom, each panel shows the valid and invalid LRP-R onset latency, and HEOG, for each response lag.

5.3.2.2.2 The LRP-R

Figure 5.16 shows the group average LRP-R and HEOG waveforms for each condition at each response tone lag. The mean onset latency of the response-locked LRP for valid and invalid trials at each response lag for Experiment 5.2 is shown in Table 5.8. A repeated-measures ANOVA on the onset latencies with Cue and Lag as factors showed a non-significant main effect of Cue ($F(1,11)=0.44$, $p>0.05$), but a significant main effect of Lag ($F(5,55)=6.45$, $p<0.001$), such that LRP-R mean onset latency increased as the response tone lag increased, with the exception of longer lags. The Cue x Lag interaction was non-significant ($F(5,55)=1.51$, $p>0.05$). The validity of the cue therefore did not differentially modulate motoric processing.

Table 5.8. The mean onset-latency (ms) of the response-locked LRP for valid and invalid trials at six response lags of Experiment 5.2 (SD in parentheses).

Lag	Valid	Invalid
30 ms	-139 (2.06)	-150 (3.54)
60 ms	-142 (2.31)	-162 (4.33)
120 ms	-152 (3.55)	-151 (4.12)
250 ms	-180 (3.60)	-164 (3.17)
500 ms	-134 (3.75)	-105 (4.77)
1000 ms	-117 (4.23)	-114 (3.85)

5.3.2.2.3 Event-related potentials

The statistical analyses of the ERP data were conducted in exactly the same way as for Experiment 5.1. Table 5.9 shows the Greenhouse-Geisser corrected F -statistics and significance levels for the main effect of Electrode, Electrode x Cue interaction, Electrode x Lag interaction, and Electrode x Cue x Lag interaction, for seventy electrodes. A main effect of Electrode was observed at all time intervals (see Table 5.9 for F -values and significance levels) except 200-260 ms ($F(69,759)=2.04$, $p>0.05$). The earliest time interval in which we can observe an Electrode x Cue interaction is the interval between 200-260 ms ($F(69,759)=4.53$, $p<0.001$), and observed again in the 260-320 ms ($F(69,759)=15.57$, $p<0.001$) and 320-380 ms ($F(69,759)=6.30$, $p<0.001$) intervals. The Electrode x Lag interaction was also significant at the later two time intervals of 260-320 ms ($F(345,3795)=5.54$, $p<0.001$) and 320-380 ms ($F(345,3795)=5.02$, $p<0.001$). There were no significant Electrode x Cue x Lag interactions at any time interval.

Table 5.9. Greenhouse-Geisser corrected F -values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data for Experiment 5.2 ($\dagger p > 0.10$, $*p < 0.05$, $***p < 0.001$).

	Time Interval						
	80-110ms	110-140ms	140-170ms	170-200ms	200-260ms	260-320ms	320-380ms
Electrode	$F=8.60$ ***	$F=8.39$ ***	$F=8.64$ ***	$F=9.94$ ***	$F=2.04$ \dagger	$F=2.57$ *	$F=3.36$ *
Electx Cue	$F=0.69$ \dagger	$F=0.60$ \dagger	$F=0.82$ \dagger	$F=1.13$ \dagger	$F=4.53$ ***	$F=15.57$ ***	$F=6.30$ ***
Electx Lag	$F=0.93$ \dagger	$F=0.82$ \dagger	$F=0.85$ \dagger	$F=0.87$ \dagger	$F=1.10$ \dagger	$F=5.54$ ***	$F=5.02$ ***
ElectxCuexLag	$F=1.18$ \dagger	$F=1.04$ \dagger	$F=1.06$ \dagger	$F=1.26$ \dagger	$F=1.00$ \dagger	$F=0.92$ \dagger	$F=0.95$ \dagger

Figure 5.17 depicts grand-averaged waveforms for valid and invalid cue conditions, collapsed across response lag, at the same parietal-occipital and occipital electrodes focussed on in Experiment 5.1. At each electrode site shown, the figure depicts an overlap in the waveforms between cueing conditions up to 200 ms post-cue presentation, indicating similar activation of the peripheral precues, eliminating the differential stimulation between the cues used in the previous experiment.

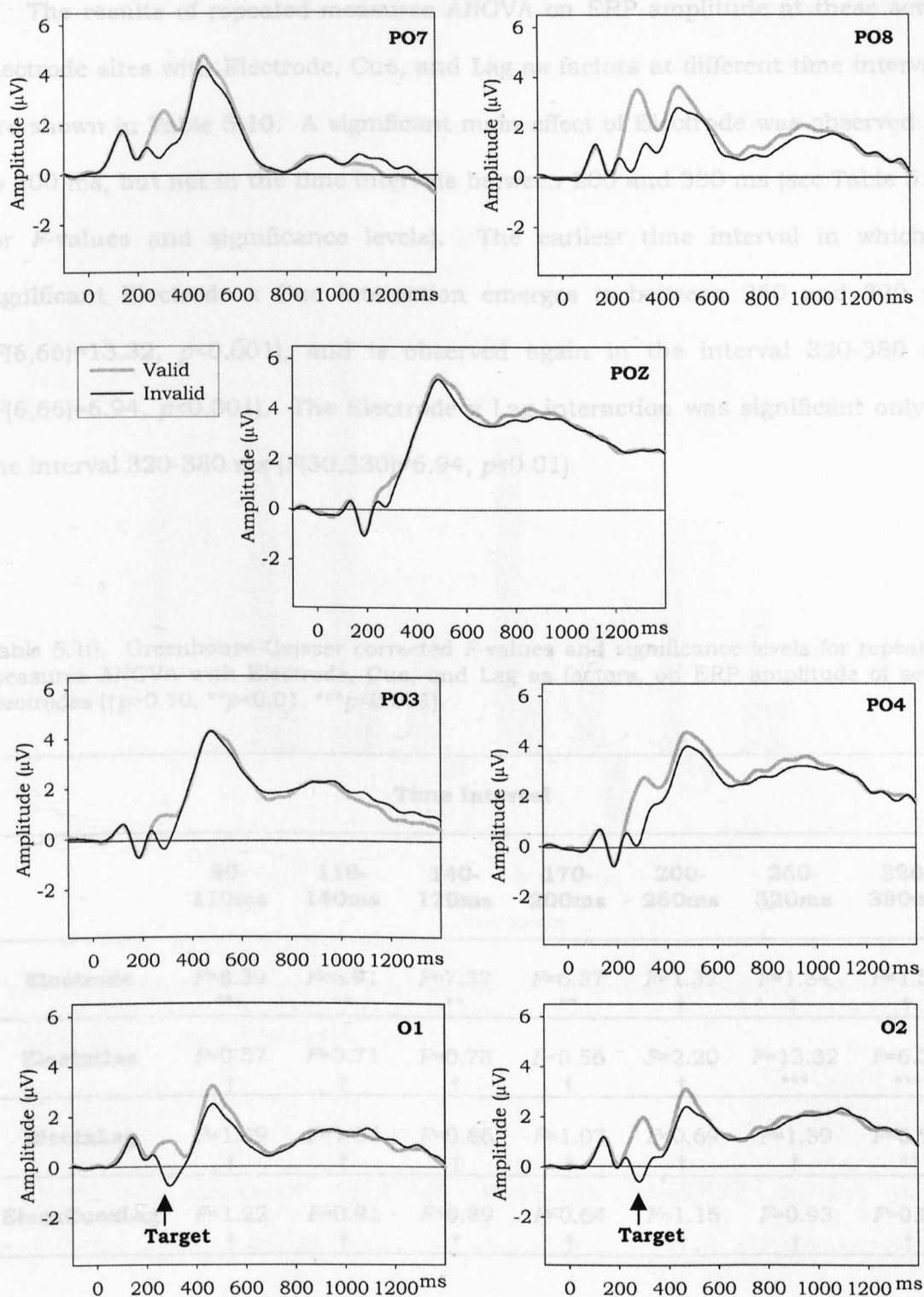


Figure 5.17. Grand-averaged waveforms for Experiment 5.2 for valid and invalid conditions. The waveforms depicted are for the same seven electrodes depicted in Figure 5.7 (PO7 PO8, POZ, PO3, PO4, O1 and O2).

The results of repeated-measures ANOVA on ERP amplitude at these seven electrode sites with Electrode, Cue, and Lag as factors at different time intervals are shown in Table 5.10. A significant main effect of Electrode was observed up to 200 ms, but not in the time intervals between 200 and 380 ms (see Table 5.10 for F -values and significance levels). The earliest time interval in which a significant Electrode \times Cue interaction emerges is between 260 and 320 ms ($F(6,66)=13.32$, $p<0.001$), and is observed again in the interval 320-380 ms ($F(6,66)=6.94$, $p<0.001$). The Electrode \times Lag interaction was significant only in the interval 320-380 ms ($F(30,330)=6.94$, $p<0.01$).

Table 5.10. Greenhouse-Geisser corrected F -values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude of seven electrodes ($\dagger p>0.10$, $**p<0.01$, $***p<0.001$).

	Time Interval						
	80- 110ms	110- 140ms	140- 170ms	170- 200ms	200- 260ms	260- 320ms	320- 380ms
Electrode	$F=8.39$ ***	$F=6.91$ **	$F=7.37$ **	$F=6.37$ **	$F=1.32$ †	$F=1.84$ †	$F=1.51$ †
Electx Cue	$F=0.57$ †	$F=0.71$ †	$F=0.78$ †	$F=0.56$ †	$F=2.20$ †	$F=13.32$ ***	$F=6.30$ ***
Electx Lag	$F=1.49$ †	$F=1.04$ †	$F=0.86$ †	$F=1.07$ †	$F=0.69$ †	$F=1.39$ †	$F=6.94$ **
ElectxCuexLag	$F=1.22$ †	$F=0.91$ †	$F=0.89$ †	$F=0.64$ †	$F=1.15$ †	$F=0.93$ †	$F=0.88$ †

Figure 5.18 depicts mean amplitude for valid and invalid trials between 260-320 ms for each electrode. Amplitude enhancements for valid trials relative to invalid trials are particularly evident at parietal-occipital sites PO7, PO8 and PO4.

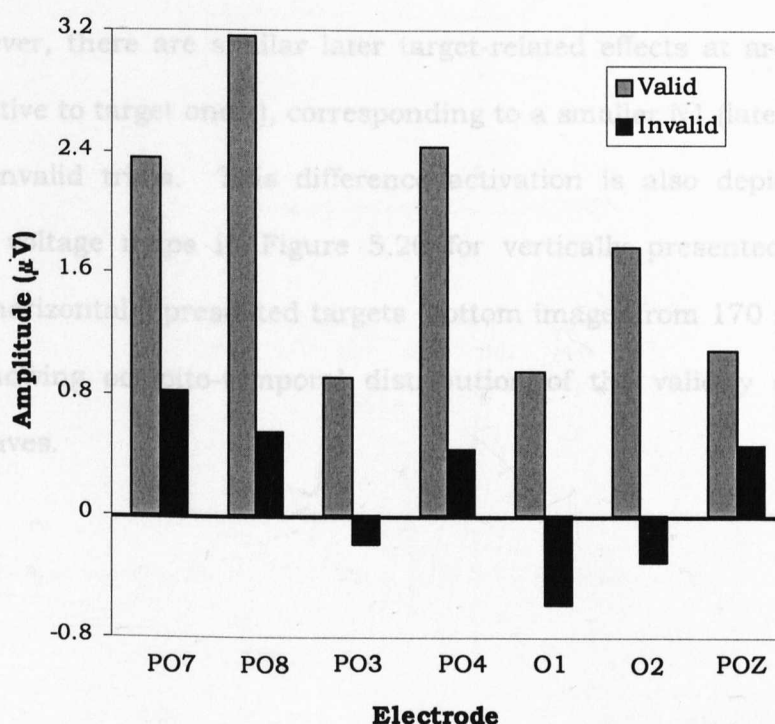


Figure 5.18. Mean amplitude for valid and invalid conditions at lateral parietal-occipital PO7 and PO8, PO3 and PO4, lateral occipital O1 and O2, and the midline parietal-occipital POZ, at time interval 260-320 ms.

Next, we calculated difference waveforms (valid(ERP activity) minus invalid(ERP activity)) to reveal the time course of when these conditions begin to differ. We calculated difference waves for when target presentation occurs along the vertical meridian at top and bottom locations, and along the horizontal

meridian, at left and right target positions. Figure 5.19 depicts these difference waveforms, and depicts the greatest difference between 250 and 300 ms, expressed as greater positivity for the valid cue relative to the invalid cue, and a bigger difference when the target was presented along the horizontal meridian than along the vertical meridian. In contrast to Experiment 5.1, there is no initial effect of the cue at 170 ms, an effect eliminated through the use of peripheral cues. However, there are similar later target-related effects at around 300 ms (180 ms relative to target onset), corresponding to a smaller N1 (latency range) for valid than invalid trials. This difference activation is also depicted in scalp topographic voltage maps in Figure 5.20 for vertically presented targets (top image) and horizontally presented targets (bottom image) from 170 ms to 450 ms post-cue, showing occipito-temporal distribution of the validity effect in ERP difference waves.

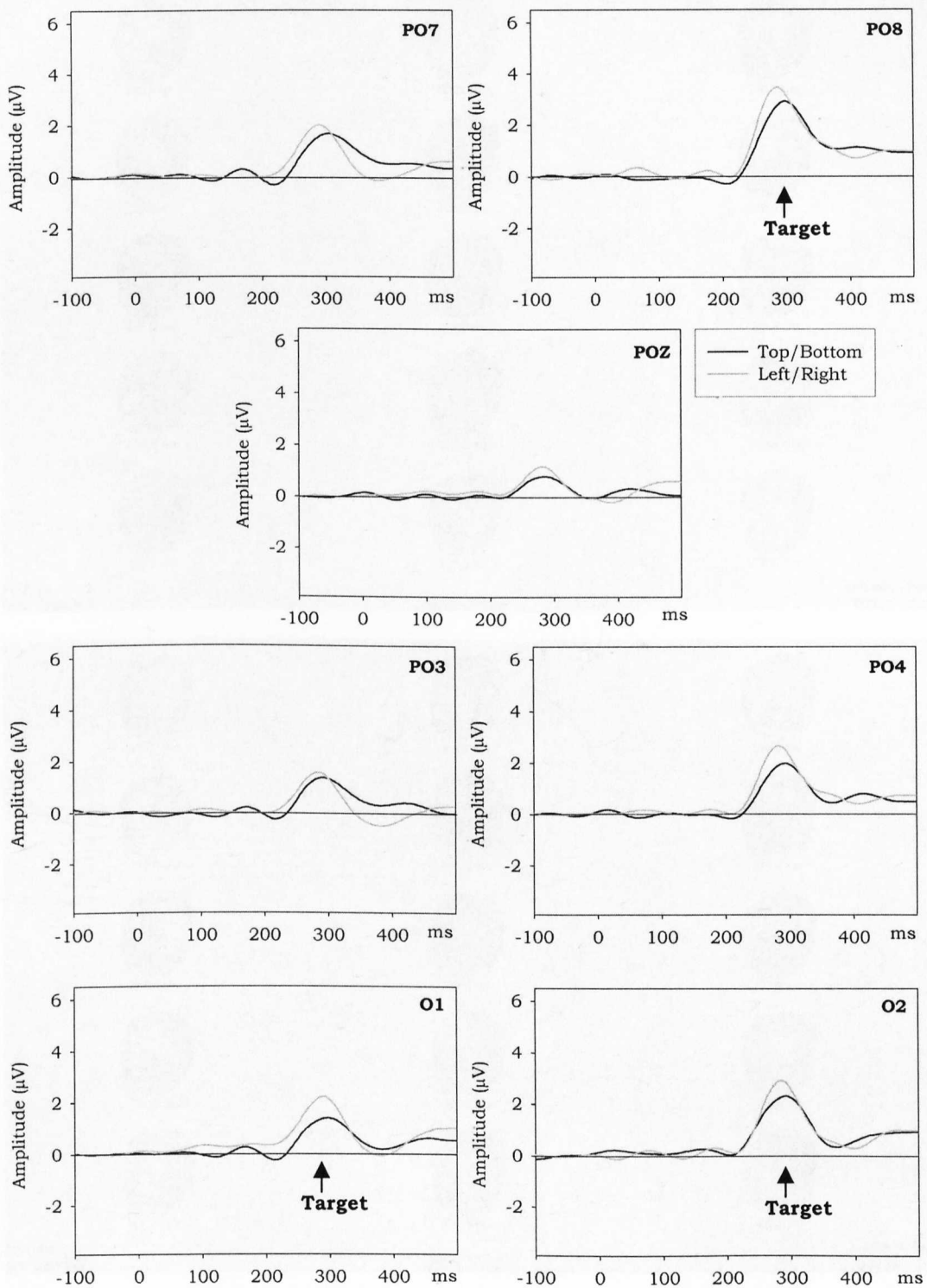


Figure 5.19. Difference waveforms between valid and invalid trials when the target stimulus was presented along the vertical meridian (collapsed across top and bottom locations) and along the horizontal meridian (collapsed across left and right target locations).

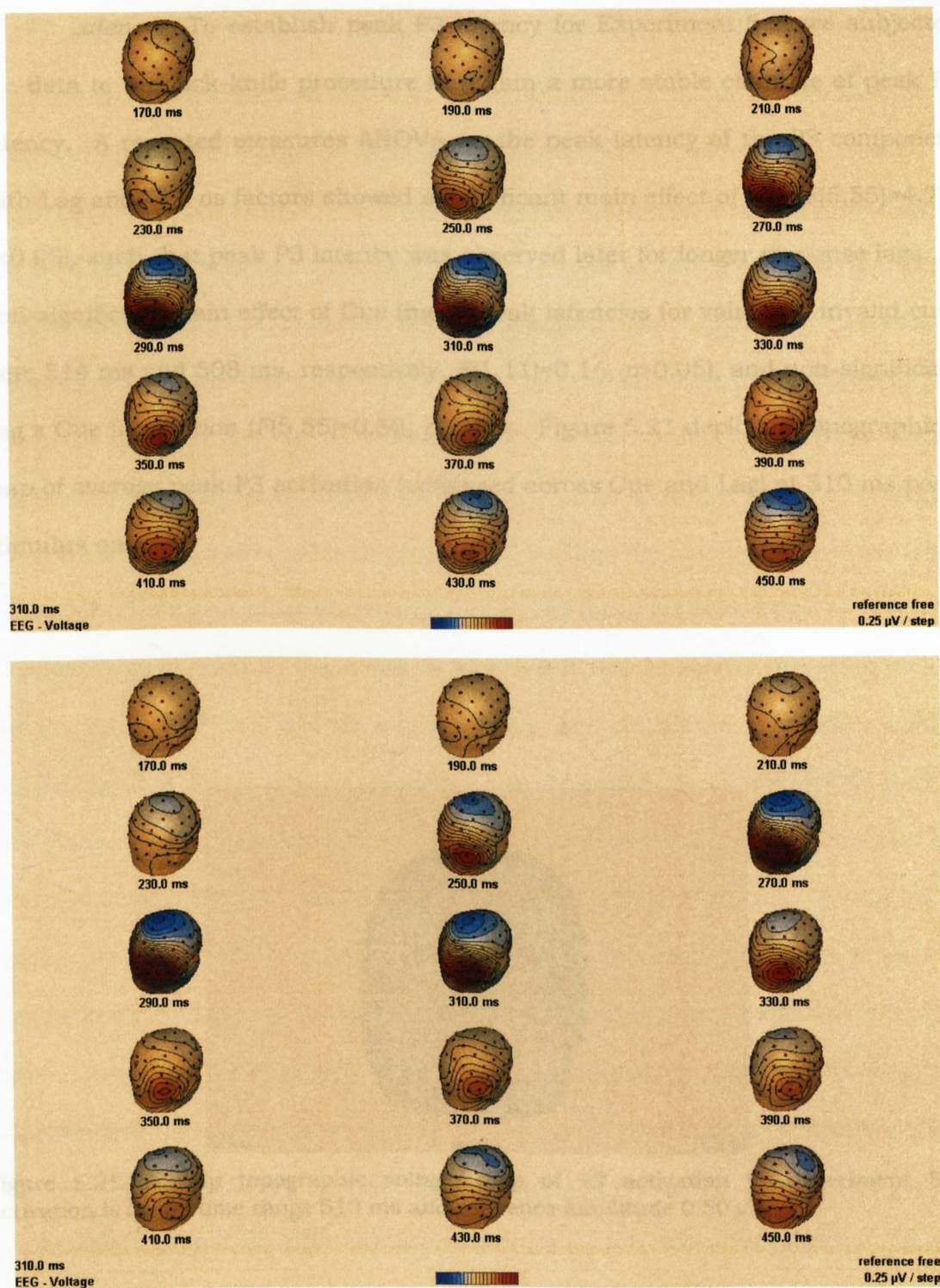


Figure 5.20. Scalp topographic voltage maps for valid minus invalid cues averaged over top and bottom target locations in the top picture and over left and right target locations in the bottom picture.

P3 Latency. To establish peak P3 latency for Experiment 5.2, we subjected the data to the jack-knife procedure to obtain a more stable estimate of peak P3 latency. A repeated measures ANOVA on the peak latency of the P3 component with Lag and Cue as factors showed a significant main effect of Lag ($F(5,55)=4.76$, $p<0.05$), such that peak P3 latency was observed later for longer response lags. A non-significant main effect of Cue (mean peak latencies for valid and invalid cues were 514 ms and 508 ms, respectively, $F(1,11)=0.16$, $p>0.05$), and non-significant Lag x Cue interaction ($F(5,55)=0.50$, $p>0.05$). Figure 5.21 depicts a topographical map of average peak P3 activation (collapsed across Cue and Lag) at 510 ms post-stimulus onset.

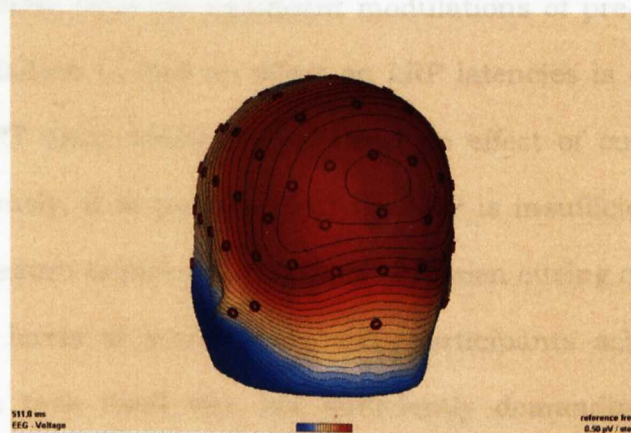


Figure 5.21. Scalp topographic voltage map of P3 activation in Experiment 5.2. Activation is in the time range 510 ms and reference amplitude 0.50 μ V.

5.3.3 Discussion

The aim of Experiment 5.2 was to determine the locus of the attentional modulation on the temporal dynamics of information processing. Similar to Experiment 5.1, the behavioural data revealed no difference in discriminability on the SAT function between valid and invalid peripheral cue conditions, indicating that attention did not enhance target discriminability. Speed of processing, however, was affected by the validity of the cue; the intercept of the SAT function occurred earlier in the validly cued condition (349 ms) compared to the invalid condition (393 ms), a difference of 44 ms. This speeding up of information processing is consistent with Experiment 5.1 where we also found a modulation in the speed of processing, also expressed on the intercept parameter.

Again, to determine the locus of this effect, we analysed the effect of cue validity on S-LRP and LRP-R onset latency and found no significant effects, suggesting that there were no significant modulations of pre-motoric or motoric processes. The failure to find an effect on LRP latencies is consistent with the analysis of the RT data, which also showed no effect of cueing. However, as mentioned previously, it is possible that the LRP is insufficiently sensitive as a chronometric measure to pick up differences between cueing conditions in such a task. The high levels of performance that participants achieved in this task suggest that the task itself was not sufficiently demanding. In contrast to Experiment 5.1, no modulation was found on the peak latency of the P3 component, indicating that stimulus categorisation was not affected by cue validity.

As for the effect of cueing on the amplitude of the ERP, the earliest modulation arose in the 260-320 ms interval, that is, 134 to 194 ms following

target onset, most evident over occipital areas, corresponding to a smaller N1 for valid trials compared to invalid trials. In relation to previous studies analysing the effect of attention on N1 amplitude, results are mixed according to the SOA used. While Fu et al. (2001) found an enhanced N1 for valid trials at short SOAs, Doallo et al. (2004) found a smaller N1 for valid trials at longer SOAs. This effect may somehow reflect an enhancement in processing for targets opposite the cued location, a finding that has been reported in a behavioural task for cues and targets that appear across the same axis, horizontal or vertical, at SOAs from 106 ms to 541 ms (Tse, Sheinberg, and Logothetis, 2003).

In a further experiment, the difficulty of the task is increased by adding distracter stimuli to the display in the hope of establishing further the performance benefits of attention, and to observe the effects on the LRP and ERPs.

5.4 Experiment 5.3

A problem with both Experiments 5.1 and 5.2 was the high level of performance attained by observers; here in Experiment 5.3 we increase the difficulty of the task by presenting the target stimulus in an array of distracters. In including distracters in the display, this experiment is more akin to the feature search condition of Carrasco and McElree (2001), in which the target stimulus is presented with three distracter stimuli. However, the pre-cueing procedure used here is different from that used by Carrasco and McElree. Similar to Experiment 5.2, peripheral valid pre-cues are used whose validity is manipulated ($p(\text{validity} = .5)$) to direct attention towards the target on half of the trials, and away from the

target on remaining trials. Additionally, in order to eliminate the possibility that any potential attentional effect could be explained simply in terms of reducing spatial uncertainty regarding the target's location, we also employed, in addition to a valid or invalid pre-cue, a report cue that was either consistent with the location of the pre-cue on valid trials, or inconsistent on invalid trials (Lu and Doshier, 2000). The task of the observer was to discriminate the orientation of the target stimulus that appeared in the location of the report cue.

5.4.1 Method

5.4.1.1 Participants

Twelve University of Glasgow students with normal, or corrected to normal vision, were paid for participating in the experiment (Mean age 22.2 years, all right-handed).

5.4.1.2 Stimuli and Procedure

As in Experiments 5.1 and 5.2, target stimuli were oriented gabor patches, tilted clockwise and anti-clockwise by 45 degrees. In order to make the task more attentionally demanding, we inserted distracter stimuli (vertically oriented gabor patches) into the three remaining locations. To eliminate spatial uncertainty, we also employed a report cue that comprised a small straight line to inform the observer of the target to be discriminated. The neutral and valid cues from Experiment 5.1 were replaced by a peripheral cue that accurately indicated the location of the upcoming target on 50% of the trials and away from the target on the remaining 50%, appearing equiprobably at any of the other 3 positions. As

before, the observer's task was to indicate the orientation of the target gabor patch according to the speed-accuracy trade-off procedure.

5.4.2 Results

5.4.2.1 Behavioural data

5.4.2.1.1. Reaction time and accuracy

Table 5.11 presents reaction times and performance accuracy for twelve observers for valid and invalid cue conditions at six response lags. A repeated measures ANOVA on the reaction time data with Cue and Lag as factors showed a significant main effect of Lag ($F(5,55)=490.19$, $p<0.001$), a significant main effect of Cue ($F(1,11)=100.61$, $p<0.001$), and significant Lag x Cue interaction ($F(5,55)=3.49$, $p<0.01$). The same analysis on the accuracy data showed significant main effects of Lag ($F(5,55)=31.73$, $p<0.001$) and Cue ($F(1,11)=19.61$, $p<0.001$), and significant Lag x Cue interaction ($F(5,55)=10.02$, $p<0.001$). The validity of the cue therefore modulated reaction time so that observers responded faster in the valid condition (M (valid and invalid: 582 ms vs. 624 ms), and improved performance accuracy (M (valid and invalid: 90% vs. 86% correct). The significant Lag x Cue interaction indicated a decreasing cueing effect in accuracy with increasing response lag, whereas the cueing effect on RT was largest for short and longer lags, and smaller for intermediate lags.

Table 5.11. Mean RT and performance accuracy for valid and invalid conditions at six response lags (*SD* in parentheses).

Lag	Valid RT	Invalid RT	Valid Acc (%)	Invalid Acc (%)
30 ms	391 (30)	434 (37)	85 (7.8)	69 (13.1)
60 ms	399 (30)	436 (33)	87 (8.4)	69 (12.4)
120 ms	424 (32)	456 (25)	90 (9.8)	74 (14.0)
250 ms	510 (20)	539 (16)	95 (5.4)	82 (11.4)
500 ms	701 (42)	742 (27)	96 (3.1)	88 (11.9)
1000 ms	1064 (109)	1136 (90)	97 (3.1)	89 (11.5)

5.4.2.1.2. *Speed-Accuracy Trade-off*

Figure 5.22 depicts the group average data and the fits for both the valid and invalid conditions. Adjusted- R^2 values for the group data, for the valid and invalid conditions, were .99 and .98 respectively, with individual values ranging from between .5 and .99 for the valid condition, and between .54 and .99 for the invalid condition.

intercept was observed for the valid condition at 331 ms as opposed to 353 ms in the invalid condition, a difference of 22 ms ($t(11)=2.72, p<0.01$).

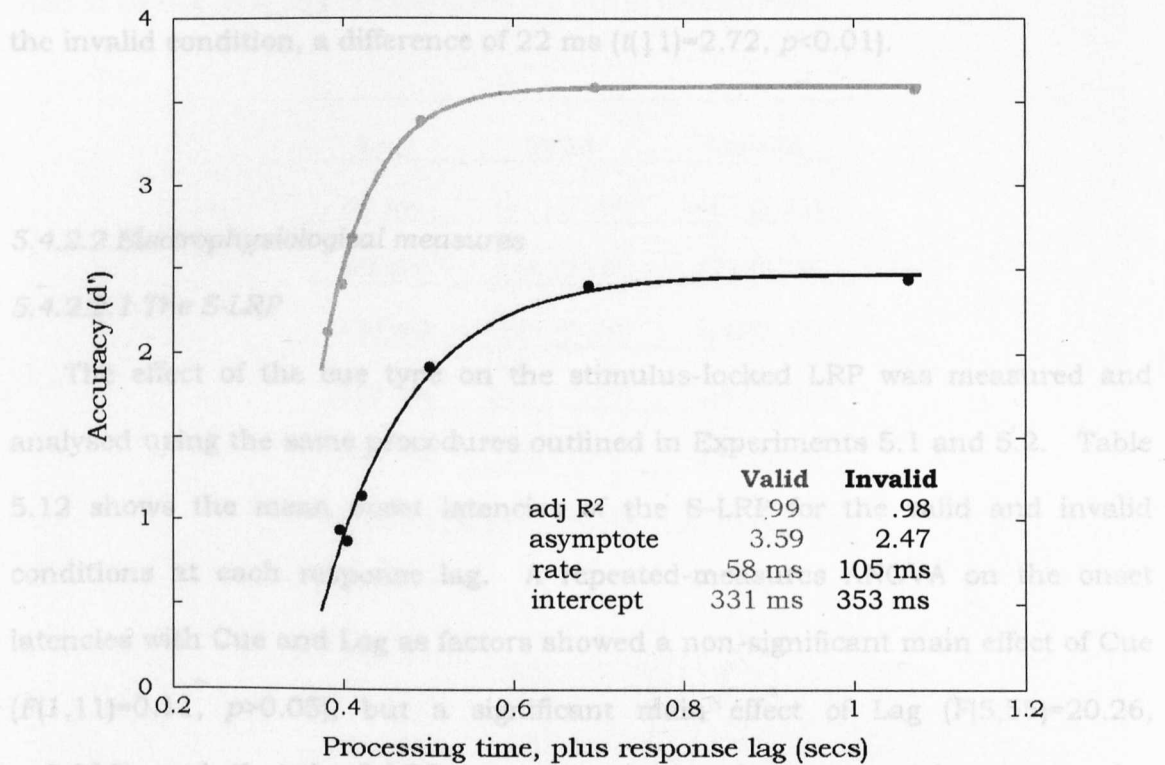


Figure 5.22. Average time course data for Experiment 5.3. Accuracy is plotted in d' units as a function of processing time (response lag plus processing time in seconds) for both valid and invalid conditions. The data is fitted with Equation 5.1.

Figure 5.22 depicts the S-LRP and NEQG waveforms for both conditions at each response lag for Experiment 5.3.

In this experiment, the effect of attention on discriminability is reflected in the difference in the asymptotic parameters of the two conditions; asymptotic performance was reached at 3.59 d' units in the valid condition and 2.47 in the invalid condition, indicating that the validity of the cue enhanced target discriminability ($t(11)=3.27, p<0.005$). To assess the effect of attention on the speed of processing, we turn to the rate and intercept parameters. Figure 5.22 shows that the rate of information processing was faster in the valid condition relative to the invalid condition (58 ms vs. 105 ms; $t(11)=2.41, p<0.01$), a difference of 47 ms. In terms of the intercept of the SAT function, an earlier

intercept was observed for the valid condition at 331 ms as opposed to 353 ms in the invalid condition, a difference of 22 ms ($t(11)=2.72$, $p<0.01$).

5.4.2.2 Electrophysiological measures

5.4.2.2.1 The S-LRP

The effect of the cue type on the stimulus-locked LRP was measured and analysed using the same procedures outlined in Experiments 5.1 and 5.2. Table 5.12 shows the mean onset latencies of the S-LRP for the valid and invalid conditions at each response lag. A repeated-measures ANOVA on the onset latencies with Cue and Lag as factors showed a non-significant main effect of Cue ($F(1,11)=0.11$, $p>0.05$), but a significant main effect of Lag ($F(5,55)=20.26$, $p<0.001$), such that the S-LRP mean onset latency increased as the response lag increased. The Cue x Lag interaction was non-significant ($F(5,55)=0.29$, $p>0.05$). Figure 5.22 depicts the S-LRP and HEOG waveforms for both conditions at each response lag for Experiment 5.3.

Table 5.12. The mean onset-latency (ms) of the stimulus-locked LRP for valid and invalid trials at six response lags of Experiment 5.3 (*SD* in parentheses).

Lag	Valid	Invalid
30 ms	481 (10.48)	467 (5.21)
60 ms	466 (2.70)	471 (8.02)
120 ms	509 (7.34)	508 (7.46)
250 ms	561 (13.94)	638 (11.69)
500 ms	898 (7.63)	857 (6.26)
1000 ms	1264 (66.48)	1153 (57.08)

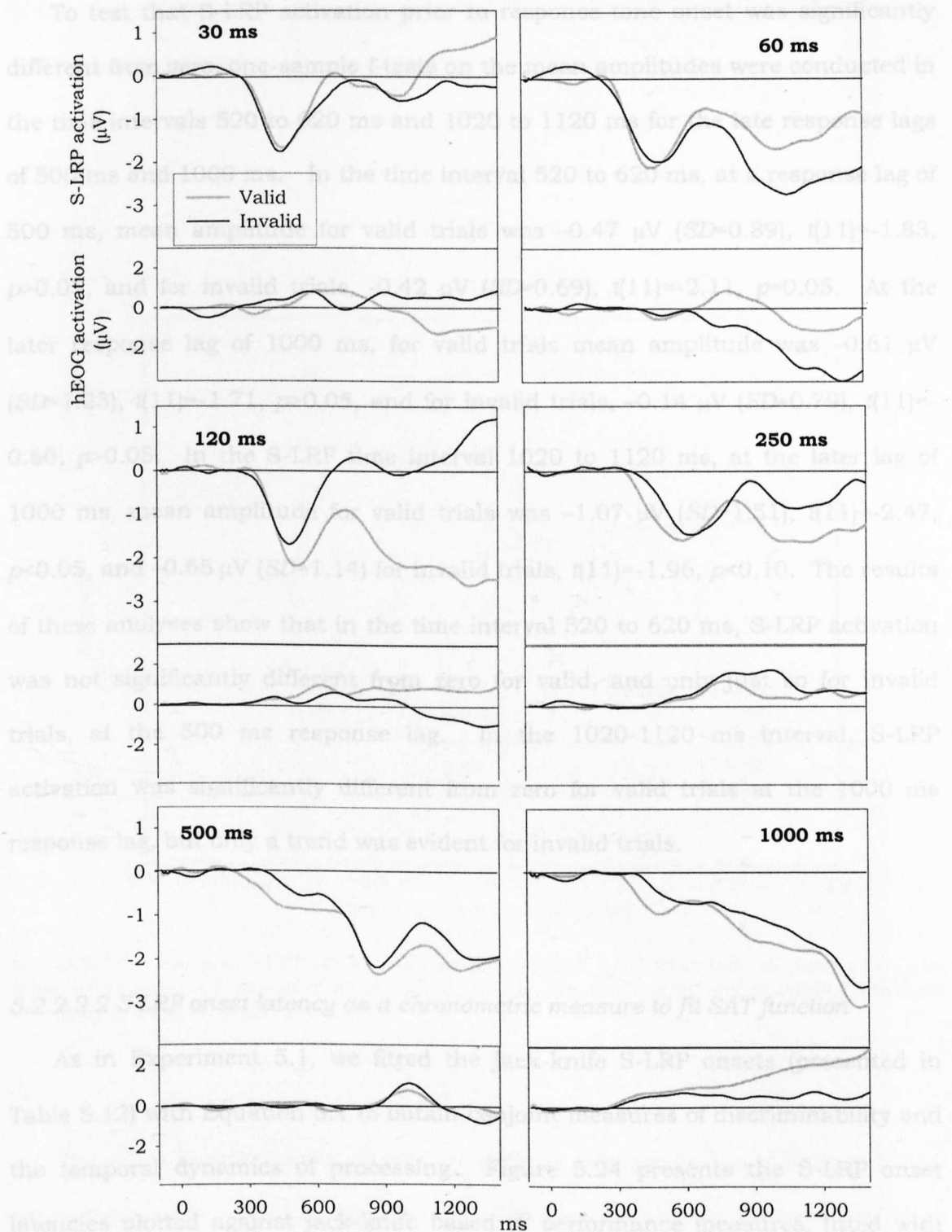


Figure 5.23. The stimulus-locked LRP and HEOG for Experiment 5.3. From left-to-right, top-to bottom, each panel shows the valid and invalid S-LRP onset latency, and HEOG, for each response lag.

To test that S-LRP activation prior to response tone onset was significantly different from zero, one-sample *t*-tests on the mean amplitudes were conducted in the time intervals 520 to 620 ms and 1020 to 1120 ms for the late response lags of 500 ms and 1000 ms. In the time interval 520 to 620 ms, at a response lag of 500 ms, mean amplitude for valid trials was $-0.47 \mu\text{V}$ ($SD=0.89$), $t(11)=-1.83$, $p>0.05$, and for invalid trials, $-0.42 \mu\text{V}$ ($SD=0.69$), $t(11)=-2.11$, $p=0.05$. At the later response lag of 1000 ms, for valid trials mean amplitude was $-0.61 \mu\text{V}$ ($SD=1.23$), $t(11)=-1.71$, $p>0.05$, and for invalid trials, $-0.14 \mu\text{V}$ ($SD=0.79$), $t(11)=-0.60$, $p>0.05$. In the S-LRP time interval 1020 to 1120 ms, at the later lag of 1000 ms, mean amplitude for valid trials was $-1.07 \mu\text{V}$ ($SD=1.51$), $t(11)=-2.47$, $p<0.05$, and $-0.65 \mu\text{V}$ ($SD=1.14$) for invalid trials, $t(11)=-1.96$, $p<0.10$. The results of these analyses show that in the time interval 520 to 620 ms, S-LRP activation was not significantly different from zero for valid, and only just so for invalid trials, at the 500 ms response lag. In the 1020-1120 ms interval, S-LRP activation was significantly different from zero for valid trials at the 1000 ms response lag, but only a trend was evident for invalid trials.

5.2.2.2.2 S-LRP onset latency as a chronometric measure to fit SAT function

As in Experiment 5.1, we fitted the jack-knife S-LRP onsets (presented in Table 5.12) with Equation 5.1 to obtain conjoint measures of discriminability and the temporal dynamics of processing. Figure 5.24 presents the S-LRP onset latencies plotted against jack-knife based d' performance measures, fitted with Equation 5.1. The quality of fit determined by the value of adjusted R^2 showed values of .87 and .99 for the valid and invalid conditions respectively. The quality

of fits of individual data ranged from .84 to .9 in the valid condition and .98 to .99 in the invalid condition.

target stimulus. To assess the effect of attention on the speed of processing, we turn again to the rate and intercept parameters. Figure 3.24 shows that there was a significant difference in the rate parameter (72 ms versus 174 ms, for valid and invalid conditions, respectively, $t(11)=10.01$, $p<0.001$), indicating that the valid cue accelerated the rate at which information was processed, a difference of 102 ms compared to the invalid cue. No significant effect of cueing was found on the intercept parameter of the SAT function, for the valid condition the intercept was 400 ms versus the slightly earlier time of 394 ms in the invalid condition ($t(11)=0.21$, $p>0.05$).

5.4.2. The LRP

Table 5.13 shows the mean parameters of the response selection LRP for valid and invalid trials at each response lag. A two-way ANOVA with Cue and Lag as factors showed the main effects of Cue and Lag to be non-significant ($F(1,11)=0.11$, $p>0.05$ and $F(1,11)=0.15$, $p>0.05$, respectively). The Cue \times Lag interaction was also non-significant ($F(1,11)=2.05$, $p>0.05$). The Cue \times Lag interaction was also non-significant

Figure 5.24. Average S-LRP onset data for Experiment 5.3. Accuracy is plotted in jack-knife d' units as a function of S-LRP onset latency for both valid and invalid conditions. The data is fitted with Equation 5.1.

both conditions at each response lag for Experiment 5.3.

The effect of attention on discriminability is reflected in the difference in the asymptotic parameters of the two conditions; asymptotic performance was reached at 3.60 d' units in the valid condition, and 2.52 in the invalid condition. A one-tailed t -test on the individual asymptote parameters was significant

($t(11)=9.92$, $p<0.001$), indicating that the validity of the cue improved the discriminability of the target stimulus. To assess the effect of attention on the speed of processing, we turn again to the rate and intercept parameters. Figure 5.24 shows that there was a significant difference in the rate parameter (72 ms versus 174 ms, for valid and invalid conditions, respectively, $t(11)=10.01$, $p<0.001$), indicating that the valid cue accelerated the rate at which information was processed, a difference of 102 ms compared to the invalid cue. No significant effect of cueing was found on the intercept parameter of the SAT function, for the valid condition the intercept was 400 ms versus the slightly earlier time of 394 ms in the invalid condition ($t(11)=0.21$, $p>0.05$).

5.4.2.2 The LRP-R

Table 5.13 shows the mean onset latencies of the response-locked LRP for valid and invalid trials at each response tone lag. A repeated-measures ANOVA with Cue and Lag as factors showed the main effects of Cue and Lag to be non-significant (main effect of Cue, $F(1,11)=2.02$, $p>0.05$, and main effect of Lag, $F(5,55)=2.05$, $p>0.05$). The Cue x Lag interaction was also non-significant ($F(5,55)=1.12$, $p>0.05$). Figure 5.25 shows the LRP-R and HEOG waveforms for both conditions at each response lag for Experiment 5.3.

Table 5.13. The mean onset-latencies (ms) of the response-locked LRP for valid and invalid trials at six response lags of Experiment 5.3 (*SD* in parentheses).

Lag	Valid	Invalid
30 ms	-146 (4.22)	-169 (5.94)
60 ms	-159 (4.20)	-160 (5.63)
120 ms	-122 (5.43)	-179 (4.12)
250 ms	-197 (4.29)	-174 (6.61)
500 ms	-129 (7.52)	-160 (5.88)
1000 ms	-130 (6.57)	-164 (6.93)

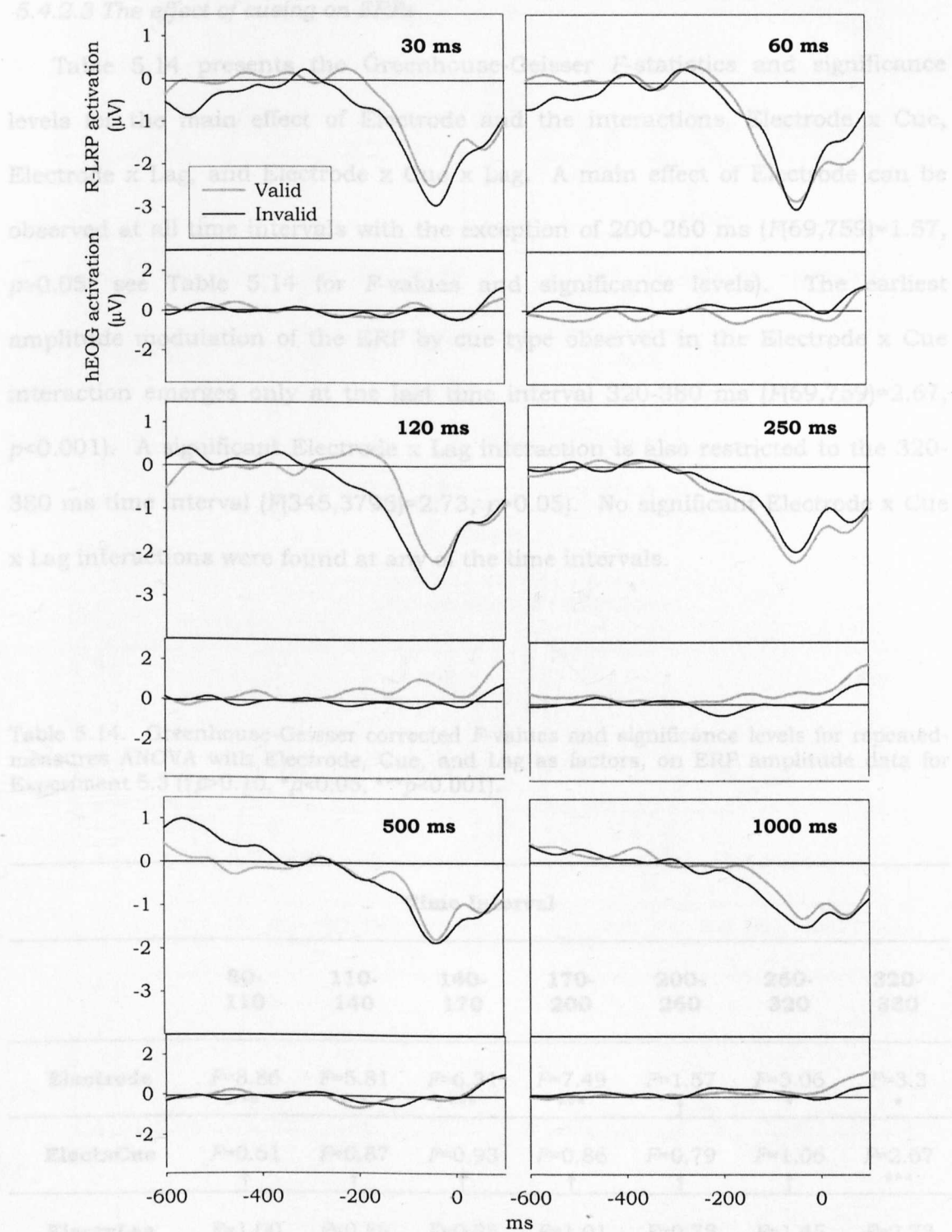


Figure 5.25. The response-locked LRP and HEOG for Experiment 5.3. From left-to-right, top-to-bottom, each panel shows the valid and invalid LRP-R onset latency, and HEOG, for each response lag.

5.4.2.3 The effect of cueing on ERPs

Table 5.14 presents the Greenhouse-Geisser F -statistics and significance levels for the main effect of Electrode and the interactions, Electrode x Cue, Electrode x Lag, and Electrode x Cue x Lag. A main effect of Electrode can be observed at all time intervals with the exception of 200-260 ms ($F(69,759)=1.57$, $p>0.05$; see Table 5.14 for F -values and significance levels). The earliest amplitude modulation of the ERP by cue type observed in the Electrode x Cue interaction emerges only at the last time interval 320-380 ms ($F(69,759)=2.67$, $p<0.001$). A significant Electrode x Lag interaction is also restricted to the 320-380 ms time interval ($F(345,3795)=2.73$, $p>0.05$). No significant Electrode x Cue x Lag interactions were found at any of the time intervals.

Table 5.14. Greenhouse-Geisser corrected F -values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data for Experiment 5.3 ($\dagger p>0.10$, $*p<0.05$, $***p<0.001$).

	Time Interval						
	80-110	110-140	140-170	170-200	200-260	260-320	320-380
Electrode	$F=8.86$ ***	$F=5.81$ ***	$F=6.31$ ***	$F=7.49$ ***	$F=1.57$ †	$F=3.06$ *	$F=3.3$ *
ElectxCue	$F=0.61$ †	$F=0.87$ †	$F=0.93$ †	$F=0.86$ †	$F=0.79$ †	$F=1.06$ †	$F=2.67$ ***
ElectxLag	$F=1.00$ †	$F=0.89$ †	$F=0.85$ †	$F=1.01$ †	$F=0.78$ †	$F=1.45$ †	$F=2.73$ ***
EleccCueLag	$F=0.99$ †	$F=0.99$ †	$F=0.97$ †	$F=0.80$ †	$F=1.07$ †	$F=0.91$ †	$F=0.77$ †

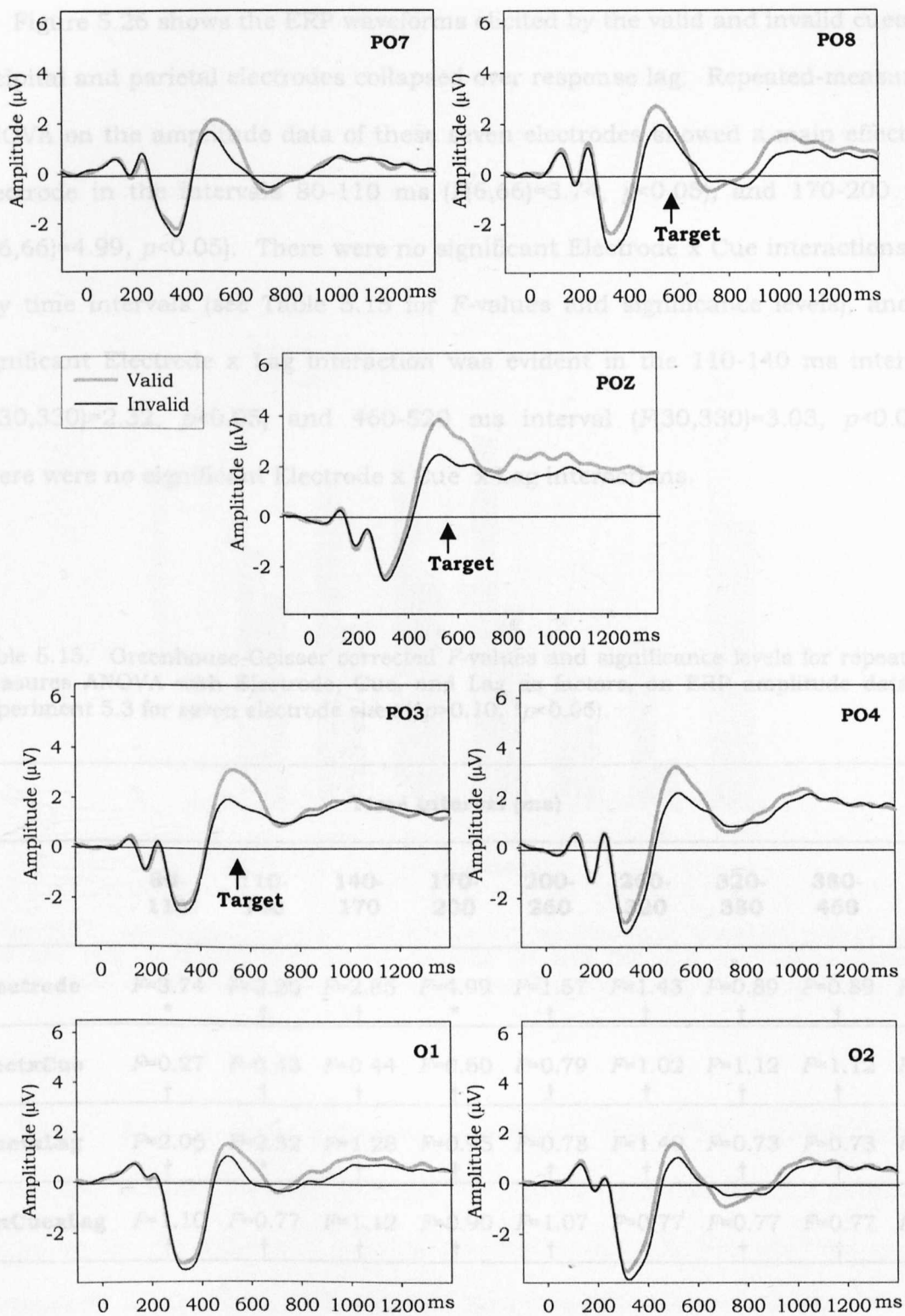


Figure 5.26. Grand-averaged waveforms for Experiment 5.3 for valid and invalid conditions at electrodes PO7 PO8, POZ, PO3, PO4, O1, and O2.

Figure 5.26 shows the ERP waveforms elicited by the valid and invalid cues at occipital and parietal electrodes collapsed over response lag. Repeated-measures ANOVA on the amplitude data of these seven electrodes showed a main effect of Electrode in the intervals 80-110 ms ($F(6,66)=3.74$, $p<0.05$), and 170-200 ms ($F(6,66)=4.99$, $p<0.05$). There were no significant Electrode x Cue interactions at any time intervals (see Table 5.15 for F -values and significance levels), and a significant Electrode x Lag interaction was evident in the 110-140 ms interval ($F(30,330)=2.32$, $p<0.05$) and 460-520 ms interval ($F(30,330)=3.03$, $p<0.05$). There were no significant Electrode x Cue x Lag interactions.

Table 5.15. Greenhouse-Geisser corrected F -values and significance levels for repeated-measures ANOVA with Electrode, Cue, and Lag as factors, on ERP amplitude data of Experiment 5.3 for seven electrode sites ($\dagger p>0.10$, $*p<0.05$).

	Time Interval (ms)								
	80-110	110-140	140-170	170-200	200-260	260-320	320-380	380-460	460-520
Electrode	$F=3.74$ *	$F=2.26$ †	$F=2.85$ †	$F=4.99$ *	$F=1.57$ †	$F=1.43$ †	$F=0.89$ †	$F=0.89$ †	$F=1.44$ †
ElectxCue	$F=0.27$ †	$F=0.43$ †	$F=0.44$ †	$F=0.60$ †	$F=0.79$ †	$F=1.02$ †	$F=1.12$ †	$F=1.12$ †	$F=2.30$ †
ElectxLag	$F=2.06$ †	$F=2.32$ *	$F=1.28$ †	$F=0.95$ †	$F=0.78$ †	$F=1.40$ †	$F=0.73$ †	$F=0.73$ †	$F=3.03$ *
ElecxCuexLag	$F=1.10$ †	$F=0.77$ †	$F=1.12$ †	$F=0.90$ †	$F=1.07$ †	$F=0.77$ †	$F=0.77$ †	$F=0.77$ †	$F=0.71$ †

Figure 5.27 depicts the mean amplitudes for valid and invalid trials between 460-520 ms (around 340-400 ms following target onset) for each electrode. Amplitude enhancements for valid trials relative to invalid trials are particularly evident at parietal-occipital sites PO8, PO4 and POZ.

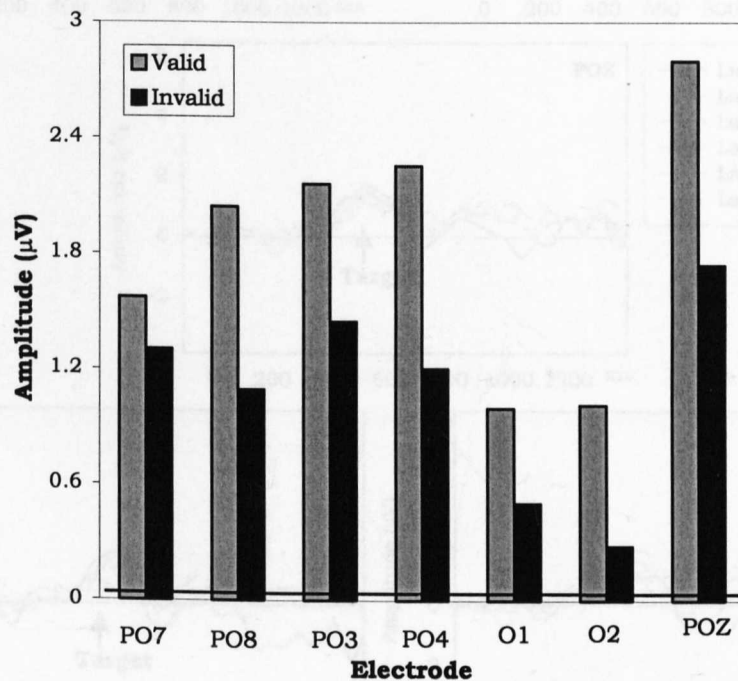


Figure 5.27. Mean amplitude for valid and invalid conditions at lateral parietal-occipital PO7 and PO8, PO3 and PO4, lateral occipital O1 and O2, and the midline parietal-occipital POZ, at time interval 460-520 ms (around 340-400 ms following target onset).

Figure 5.28 depicts difference waveforms (valid(ERP activity) minus invalid(ERP activity)) for each response lag. The figure shows that the pattern of activation for the valid cue is more positive in the 400-600 ms time range, particularly at electrodes PO3 and Pz. Scalp topographic voltage maps of this difference activation are shown in Figure 5.29, confirming the occipito-temporo-parietal distribution of the validity effect.

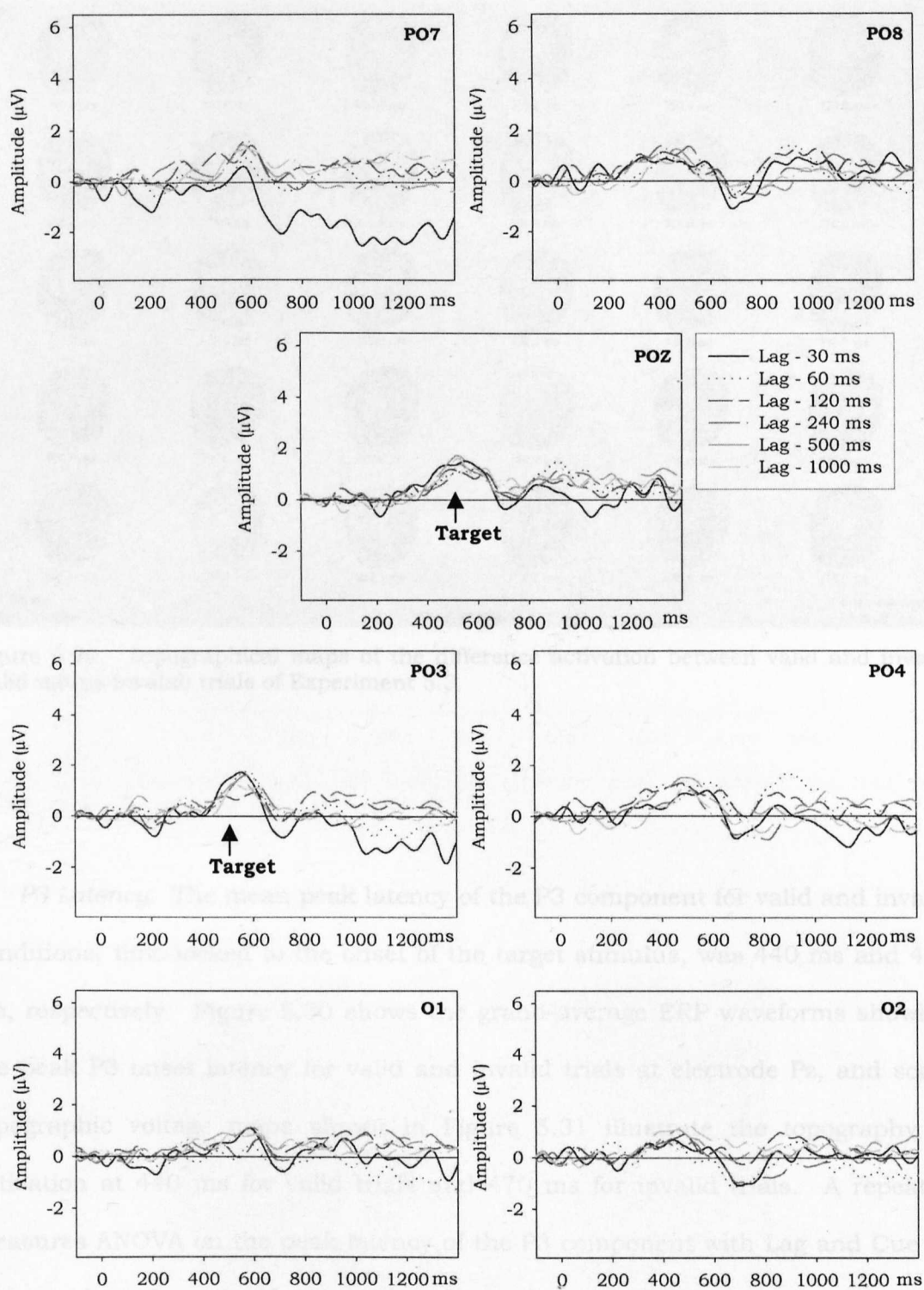


Figure 5.28. Difference waves (valid(ERP) minus invalid(ERP)) for the six response lags of Experiment 5.3.



Figure 5.29. Topographical maps of the difference activation between valid and invalid (valid minus invalid) trials of Experiment 5.3.

P3 Latency. The mean peak latency of the P3 component for valid and invalid conditions, time-locked to the onset of the target stimulus, was 440 ms and 470 ms, respectively. Figure 5.30 shows the grand-average ERP waveforms showing the peak P3 onset latency for valid and invalid trials at electrode Pz, and scalp topographic voltage maps shown in Figure 5.31 illustrate the topography of activation at 440 ms for valid trials and 470 ms for invalid trials. A repeated measures ANOVA on the peak latency of the P3 component with Lag and Cue as factors showed a significant main effect of Lag ($F(5,55)=13.51$, $p<0.001$), a significant main effect of Cue ($F(1,11)=7.68$, $p<0.05$), and non-significant Lag x Cue interaction ($F(5,55)=0.39$, $p>0.05$). The validity of the cue significantly

modulated the latency of the peak P3 component, indicating an effect at the level of stimulus categorisation.

Experiment 5.3 can be summarised as following. Firstly, cue validity enhanced target discriminability as evidenced in the higher asymptote parameter of the SAT function. Secondly, cue validity speeded up

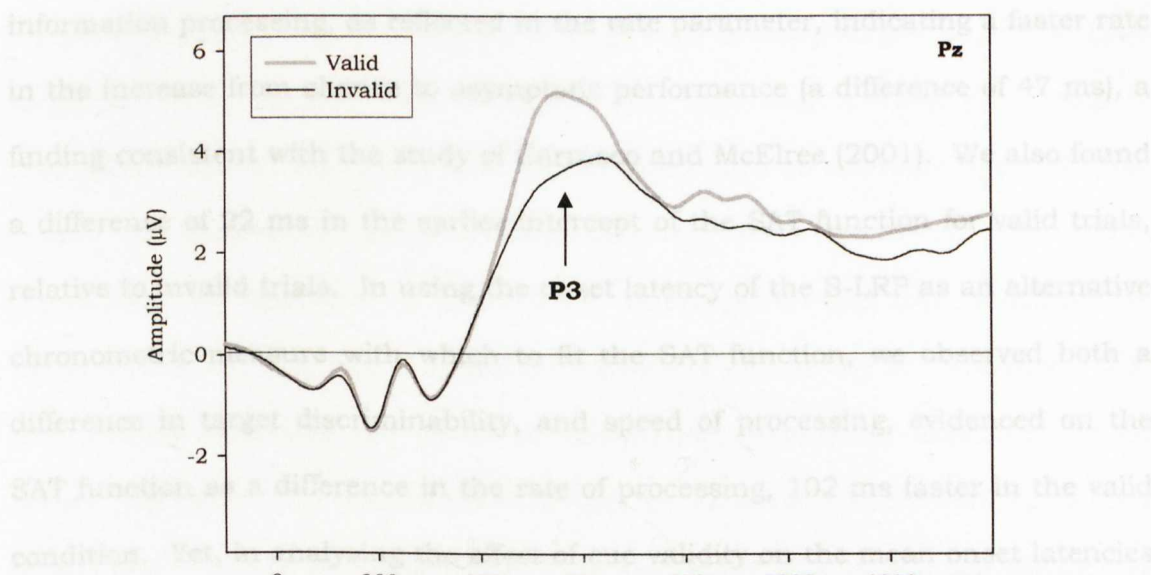


Figure 5.30. Grand-averaged ERP waveform showing peak P3 latency for valid and invalid trials at electrode Pz.

any effect, indicating that advance information regarding target location did not shorten pre-motor or motor processing, as measured by the LRP.

Turning up the ERP data, the earliest target-related amplitude modulation

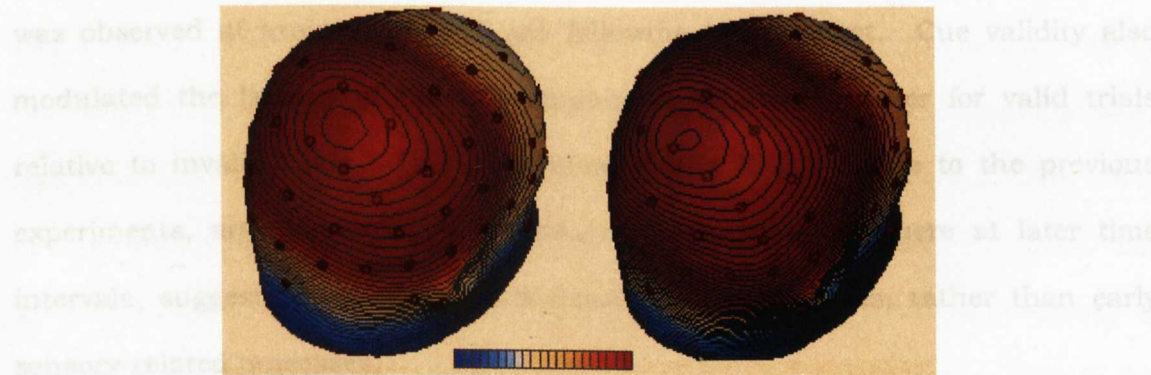


Figure 5.31. Scalp topographic voltage maps of P3 peak latency for valid (left) and invalid (right) trials, at 440 ms and 470 ms, respectively, for Experiment 5.3.

5.4.3 Discussion

The main results of Experiment 5.3 can be summarised as following. Firstly, cue validity enhanced target discriminability as evidenced in the higher asymptote parameter of the SAT function. Secondly, cue validity speeded up information processing, as reflected in the rate parameter, indicating a faster rate in the increase from chance to asymptotic performance (a difference of 47 ms), a finding consistent with the study of Carrasco and McElree (2001). We also found a difference of 22 ms in the earlier intercept of the SAT function for valid trials, relative to invalid trials. In using the onset latency of the S-LRP as an alternative chronometric measure with which to fit the SAT function, we observed both a difference in target discriminability, and speed of processing, evidenced on the SAT function as a difference in the rate of processing, 102 ms faster in the valid condition. Yet, in analysing the effect of cue validity on the mean onset latencies of the S-LRP and LRP-R, similar to the previous two experiments, we failed to find any effects, indicating that advance information regarding target location did not shorten pre-motoric or motoric processing, as measured by the LRP.

Turning to the ERP data, the earliest target-related amplitude modulation was observed at around 300-400 ms following target onset. Cue validity also modulated the latency of the P3 component, occurring earlier for valid trials relative to invalid trials. The modulation of the ERP, relative to the previous experiments, and compared to previous studies, manifests here at later time intervals, suggesting an effect on decision-related processes, rather than early sensory-related processes.

5.5 Concluding remarks

The aims of this study were to determine if attention could speed up visual information processing, and if so, to establish the locus of such effects within the information processing stream. Potentially, the locus of a modulation on the temporal dynamics of processing could lie at early sensory stages, and/or at later response-related stages. The results of three experiments using pre-cueing to manipulate spatial attention, and a SAT procedure to examine the full time course of processing, indicated that attention does speed up information processing, as observed in the earlier intercept of the SAT functions of Experiments 5.1, 5.2, and 5.3, and in the rate of processing in Experiment 5.3. Cue validity enhanced the speed of information processing, expressed as an earlier departure from chance performance, by 22 ms in Experiment 5.3, 32 ms in 5.1, and 44 ms in Experiment 5.2, while the rate of information processing was accelerated by 47 ms in Experiment 5.3.

To determine the locus of this effect, we analysed the early pre-motoric, and later motoric, intervals of the LRP. Perhaps due to insufficient sensitivity of the LRP as a chronometric measure, no effect of cue validity was observed on either LRP interval. ERP analyses, however, revealed that the amplitude and latency of the early P1 component was unaffected by attention, while a larger N1 component was evident for invalid trials relative to valid trials in Experiment 5.2. The effect of cue validity on ERP amplitude appeared later in Experiment 5.3, occurring around 300-400 ms following target onset. While amplitude enhancements of ERP components have been widely documented, and behavioural measures of the temporal dynamics of processing, such as RT and SAT, suggest that attention speeds up processing, no equivalent effect has been reported in terms of a latency

shift for early ERP components. The results reported here are consistent with previous studies; attention did not affect the latency of early ERP components, but did incur earlier peak latency for the later P3 component. Taken together, analyses of the ERPs suggest that there were no attentional modulations of early sensory-related components, neither in amplitude nor latency. Rather, ERP amplitude was modulated at later time intervals, and the earlier latency for the peak P3 component for valid trials indicates an effect of attention at later decision-related stages related to stimulus categorisation.

Further exploration in establishing the locus of a temporal modulation in information processing, as a result of attention, should perhaps employ a more difficult task, such as a feature conjunction task, as performance levels were already high here at short response lags with invalid cues. Increasing the difficulty of the task, and thus reducing performance levels overall, may produce a more pronounced effect on the LRP and ERP waveforms.

Chapter 6

General Discussion

The aim of this thesis was to explore different questions regarding the selective use of visual information for perception and recognition. To this end, a series of experiments were conducted, each described in Chapters 2 to 5. In this final chapter, summaries of those experiments are provided, followed by a consideration of methodological and theoretical implications.

6.1 Summary

In Chapter 2, I posed the question: What is the information underlying the stable percepts of a bi-stable image? I addressed this question in Experiment 2.1 using the bubbles method of Gosselin and Schyns (2001), to determine the information underlying the bi-stable percepts of Dali's *Slave Market with the Disappearing Bust of Voltaire*. The results of this experiment showed that the information underlying these percepts differed across the 2D image and across the third dimension of spatial scales; the perception of the Nuns was restricted to

high spatial frequency information, whereas the information underlying the perception of Voltaire was constrained to lower spatial frequencies. I proposed that the alternation between the two different percepts of this bi-stable image might be explained in terms of a switch between the spatial filters processing the image.

In Experiment 2.2, I validated this information using frequency-specific adaptation in which observers adapted either to high or low spatial frequency patterns, and then viewed a hybrid image containing the information underlying the perception of the Nuns and of Voltaire (derived from Experiment 2.1). Adapting observers to high spatial frequency patterns suppressed perception of the Nuns, and induced a perception of Voltaire, whereas adaptation to low spatial frequency patterns suppressed the Voltaire percept, inducing a perception of the Nuns. Experiment 2.2 demonstrates that adaptation to the spatial frequency information underlying a percept caused a suppression of the percept, inducing a perception of the alternative image interpretation. I concluded from these two experiments that knowing the information content of a stimulus can contribute to our understanding of the perceptual mechanisms engaged in processing visual input.

In Chapter 3, the adaptation method employed in Chapter 2 was further developed to suppress the use of *local* stimulus information. To test the potential effectiveness of this method in suppressing local perceptual information and thus test the relevancy of this information for perception and recognition, in Experiment 3.1 I extended the findings of Experiment 2.2 and adapted observers locally to the information underlying the stable percepts of Dali's painting. Similar to Experiment 2.1, I adapted observers to the low spatial frequency

information underlying the Voltaire percept or to the high spatial frequency information underlying the Nuns percept. The adapting patterns however were restricted either to the spatial region corresponding to the percept (Congruent conditions), or were inconsistent with the percept (Incongruent conditions). On presentation of a hybrid image that comprised both percepts, the distribution of responses in the Congruent conditions were consistent with the findings of Experiment 2.2, that is, observers adapting to high spatial frequency information constrained to the spatial regions of the Nuns percept reported a perception orthogonal to the adapting pattern, in this case Voltaire. Participants who adapted to low spatial frequency information restricted to the spatial regions of Voltaire also reported a perception orthogonal to the adapting pattern. In contrast, the distribution of responses for the Incongruent conditions differed from the Congruent conditions; most observers who adapted to low spatial frequency information constrained to the spatial regions of the Nuns percept still perceived Voltaire, indicating a failure to suppress the Voltaire percept. The distribution of responses for observers who adapted to high spatial frequency information restricted to the spatial regions of the Nuns showed an orthogonal perception, that is they perceived Voltaire, indicating a suppression of the Nuns percept, albeit less than in the Congruent condition. The results of Experiment 3.1 highlight the importance of the interaction between the spatial regions and spatial frequencies in driving these percepts. Methodologically, I suggested that this adaptation effect on local stimulus information might provide a tool to examine the use of local stimulus information in mediating perception and recognition of visual input.

Experiment 3.2 applied the local adaptation method to the perception of gender in faces. I created hybrid face stimuli that contained the information of one face across the facial regions and spatial scales diagnostic to resolve a gender task, and the face of another gender in the non-diagnostic regions. Observers adapted to the average spatial frequency information underlying the faces, restricted to either the diagnostic regions, or to the non-diagnostic regions. It was expected that this local adaptation would result in a suppression of the visual information in the adapting region, so that gender judgements of hybrid faces would result in a distribution of responses that would reflect the face represented in the non-adapting region. The results, however, failed to find an effect of the region of adaptation on the gender judgement of the hybrid face stimuli. I discussed this null result in Chapter 3 as a methodological problem, most probably with the hybrid faces used, and in our conception of *diagnostic* information.

In Chapter 4, I addressed the following questions relating to the evolution of information use with learning: How does attending to visual information evolve over time as a result of learning? And what is the mechanism(s) enabling learning? Since observers use only a subset of the information available, do the mechanism(s) enabling learning and selection change over the visual input? I attempted to answer these questions in a face discrimination task using the method of noise masking in a perceptual learning paradigm. The method involved measuring observers contrast thresholds in different levels of external noise across different regions of the face in several learning sessions. To characterise the mechanisms enabling learning, contrast thresholds were plotted as a function of external noise density for each day of learning, as conducted

previously by Gold et al. (1999b, 2004). Each potential mechanism yields a different performance signature on contrast versus noise functions. To obtain a benchmark of performance we also presented the same task to a theoretically ideal observer that can use all of the information available in the stimulus. Comparison of human performance to that of ideal performance enabled the calculation of a measure of the efficiency with which human observers used the information available in each facial region, on each learning session.

The results for two observers were consistent showing increases in contrast thresholds across facial regions as noise density increased. Contrast thresholds also decreased with learning, across all image regions and four noise density levels for the same two observers. For each of the three observers tested, at some noise levels and face regions, it was not possible to establish a contrast threshold that was within the tested range, leading to several empty data cells. Therefore, with incomplete data, I was unable to fit a model to the contrast versus noise functions, such as the LAM used by Gold et al. (1999b, 2004) or the PTM of Lu and Doshier (2004), to establish the mechanisms underlying learning. The results of a third observer were inconsistent, showing decreases in efficiency for some facial regions between the first day of learning and the second and third days, before increasing again on the final session. I explained the inconsistency in this observer's results in relation to the strategies the observer may used to resolve the task.

My conclusions from Experiment 4.1 while limited by the disappointing missing data points, endorse the endeavour to provide a qualitative as well as a quantitative account of perceptual learning in a face discrimination task. The

experiment was discussed in terms of methodological improvements for future study.

Chapters 2 to 4 each investigated the visual information attended for different perceptual and recognition tasks, in Chapters 2 and 3 for the perception of a bistable image, as well as the perception of gender in faces in Chapter 3, to the evolution of information use across time in a face discrimination task in Chapter 4. Regarding the role of visual information use for perception and recognition, Chapter 5 adopted a different approach. Here, rather than looking at the information content of a stimulus I posed the question: Does attention speed up information processing? And if so, what is the locus of the attentional effect? Experiments 5.1, 5.2, and 5.3 all examined this question using a spatial attention paradigm, in which attention was manipulated using pre-cues, and speed of processing measured using SAT procedures. The locus of the effect was established by recording stimulus and response-related LRP and components of the ERPs.

The results from Experiments 5.1, 5.2 and 5.3 showed that attention does enhance the speed of visual information processing, evident in the earlier intercept (all three experiments) and rate parameters (Experiment 5.3) of the SAT function for attended versus unattended trials. In relation to the locus of this effect, analyses of the LRP revealed no effect of attention on the early stimulus-locked component, or on the late response-locked LRP stage, perhaps due to the insufficient sensitivity of the LRP as a chronometric measure. As for the event-related brain potentials, no amplitude or latency modulations were evident on early sensory-related P1/N1 components. An effect on the peak latency of the P3 component was observed, however, in Experiments 5.1 and 5.3, suggesting an

acceleration of categorisation processes. Overall, the results are consistent with the idea that attention modulates decision-related processes. For example, Schneider and Bavelier (2003) arrived at a similar conclusion in a study investigating attentional modulations on the temporal dynamics of information processing.

6.2 Issues, methodological and theoretical

Here, I consider specific methodological and general theoretical issues related to the experiments just described. Firstly, I consider the effectiveness of the bubbles method as a tool in which to determine the subset of information an observer selects to resolve perceptual and recognition tasks. Secondly, I contemplate the concept of *diagnosticity* and suggest its interpretation here may be problematic in at least one experiment. Lastly, I reflect on the form of question: What is the information selected? I suggest in many areas of research this question is of importance, but that it has its limitations.

6.2.1 Does *Bubbles* induce atypical observer strategies?

Experiments 2.1 and 2.2 together provide a case study on the visual information underlying the perception of a bi-stable image. To determine this information, I used the bubbles method of Gosselin and Schyns (2001), which was applied across the 2D image, and across different spatial scales, fine-to-coarse. The Bubbles method is unique in enabling a simultaneous search of the 2D spatial information underlying a percept, and the entire spatial frequency

spectrum decomposed into several spatial frequency bands. Other techniques such as noise masking or reverse correlation, at this time in their development, do not enable such a large search space to be sampled simultaneously for information; rather these techniques typically measure performance in stimuli band-passed to contain only a range of spatial frequencies as opposed to the entire image spectrum. The effectiveness of the bubbles method as a tool to study information use, however, has recently been subject to some criticism (Murray and Gold, 2004a, b).

At the core of this debate is the critique that in sparsely sampling a stimulus using a technique such as bubbles, observers employ strategies atypical of everyday perception and recognition. Yet, the evidence from studies applying the bubbles method to questions of information use in face categorisation tasks converge well with similar studies using different methodologies. For example, in the experiments of Gosselin and Schyns (2001) and Schyns et al. (2002), the face information used to resolve a face identity task, determined using bubbles, is highly similar to the facial information fixated in experiments recording eye movements (Yarbus, 1967). Additionally, the information used to resolve the gender task in these bubbles experiments revealed a bias to information on the left side of the face, a result consistent with studies using chimeric face stimuli (Burt and Perrett, 1997; Butler et al., 2005), and with results using reverse correlation (Mangini and Biederman, 2004). This convergence of evidence for the information observers use in different face tasks, derived from different methodologies, suggests that the sampling method of bubbles does not lead observers to adopt an atypical selective use of information. In the context of the current experiment, the technique was applied to determine the information

present when an observer experienced one percept as opposed to the other. This information was validated in Experiment 2.2 using frequency-specific adaptation, and observed orthogonality in responses suggested that the information derived using bubbles is indeed the information that underlies these percepts.

6.2.2 The concept of diagnosticity

The concept of diagnosticity was introduced here within the context of the Diagnostic Recognition Framework (Schyns, 1998). Within this framework, the information diagnostic to resolve a task is determined by the constraints imposed by the demands of the task, and the availability of stimulus information. Thus, this framework posits a *flexible* approach towards the selective use of information. Previous research using the bubbles method substantiates this approach, demonstrating that the diagnostic information in the same stimulus set differed as a function of the categorisation task demands (Gosselin and Schyns, 2001; Schyns et al., 2002). The application of the bubbles method to determine the diagnostic information to resolve a task operates by assigning a diagnostic status to a feature (at a pixel level) as it reaches a pre-determined threshold criterion, those features that fall below this criterion are assigned a non-diagnostic status, whereas those above are heralded diagnostic. The problem with this binary representation of feature diagnosticity is that in reality, it is more likely that features are best represented as a gradient of diagnosticity in which each feature has a differential diagnostic weight. At different levels of performance criterion, for example adjusting from 75% to 60% correct, a different set of features will be deemed diagnostic, however overlapping. A distinction is

necessary, therefore, between the sufficiency versus necessity of a feature, to resolve a given perceptual or recognition task.

6.2.3 Determining the information selected for perception

Here, I consider the usefulness of asking the question: What is the information selected? The flexible approach to the selective use of information states that information selection is determined by task demands, the availability of stimulus information, and previous experience with the input. Therefore, the information selected to resolve a task will vary accordingly. It could be argued that there are clear limitations in a research agenda that aims to determine solely the information subtending different perceptual or recognition tasks. While there are many research domains where a determination of the subset of stimulus information that is used to resolve a task would prove fruitful, a research agenda whose sole endeavour is to establish such information is limited where there is an absence of interest for characterising the mechanisms of processing. This is an important point to acknowledge, because the development of new techniques, such as bubbles and reverse correlation, which pictorially reveal the information, used or represented, to resolve a task, may lead to a less-than-informative application of these techniques to questions of visual processing. However, used appropriately, and in conjunction with other techniques, these techniques have the potential to provide a fuller account of visual and cognitive processing that was previously unavailable.

6.3 Conclusion

In summary, this thesis has addressed different aspects of attention to visual information for perception and recognition, including an analysis into the visual information underlying the perception of an ambiguous image, the tracing of the evolution of information use as a result of learning, and establishing the locus of an attentional modulation on the speed of visual information processing. On reflection, the main contributions of the present research are, firstly, in highlighting the importance of understanding the information content of a stimulus and how this may influence processing, and secondly, in developing methods in which to determine and validate this information. Lastly, it was shown that attention accelerates the temporal dynamics of processing, at the locus of a decision stage in the stream of information processing, related to stimulus categorisation. In addressing each research question, different methodologies were employed, the limitations of which have been discussed. In addition to highlighting methodological weaknesses of the present experiments, I have also suggested improvements to overcome these in future research, and highlighted some theoretical issues implied by the results reported here.

References

- Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P.G., & Damasio, A.R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, **433**, 68-72.
- Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences, USA*, **90**, 5718-5722.
- Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, **387**, 401-406.
- Ahumada, A.J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision*, **2**, 121-131.
- Attneave, F. (1971). Multistability and Perception. *Scientific American*, **225**, 62-71.
- Baldassi, S., Burr, D., Carrasco, M., Eckstein, M., & Verghese, P. (2004). Visual attention. *Vision Research*, **44**, 1189-1191.
- Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, **27**, 953-965.
- Barlow, H.B. (1980). The absolute efficiency of perceptual decisions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **290**, 71-82.

- Bashinski, H.S., & Bacharach, V.R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception & Psychophysics*, **28**, 241-248.
- Beard, B.L., & Ahumada, A.J. (1998). A technique to extract relevant image features for visual tasks. In B.E. Rogowitz & T.N. Pappas, (Eds.), *Human Vision and Electronic Imaging III, SPIE Proceedings*, **3299**, 79-85.
- Bennett, P.J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological Science*, **12**, 76-80.
- Biederman, I., & Gerhardstein, P.C. (1995). Viewpoint-dependent mechanisms in visual object recognition: a critical analysis. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 1506-1514.
- Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, **96**, 145-167.
- Blake, R., & Logothetis, N. (2001). Visual completion. *Nature Neuroscience Reviews*, **3**, 1-11.
- Blakemore, C., & Campbell, F.W. (1969a). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, **203**, 237-260.
- Blakemore, C., & Campbell, F.W. (1969b). Adaptation to spatial stimuli. *Journal of Physiology*, **200**, 11-13.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.
- Burgess, A.E., & Colborne, B. (1988). Visual signal detection. IV. Observer inconsistency. *Journal of the Optical Society of America*, **5**, 617-627.

- Butler, S., Gilchrist, I.D., Burt, D.M., Perrett, D.I., Jones, E., & Harvey, M. (2005). Are the perceptual biases found in chimeric face processing reflected in eye-movement patterns? *Neuropsychologia*, **43**, 52-59.
- Carrasco, M., Giordano, A.M., & McElree, B. (2004). Temporal performance fields: visual and attentional factors. *Vision Research*, **44**, 1351-1365.
- Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5363-5367.
- Carrasco, M., McElree, B., Denisova, K., & Giordano, A.M. (2003). Speed of visual processing increases with eccentricity. *Nature Neuroscience*, **6**, 699-700.
- Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research*, **40**, 1203-1215.
- Carrasco, M., Williams, P.E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: Support for signal enhancement. *Journal of Vision*, **2**, 467-479, <http://journalofvision.org/2/6/4/DOI10.1167/2.6.4>.
- Cheal, M., & Lyon, D.R. (1991). Central and peripheral precuing of forced-choice discrimination. *The Quarterly Journal of Experimental Psychology*, **43A**, 859-880.
- Coles, M.G.H. (1989). Modern mind-brain reading: Psychophysiology, physiology, and cognition. *Psychophysiology*, **26**, 251-269.
- Coles, M.G.H., Gratton, G., & Fabiani, M. (1990). Event-related brain potentials. In J.T. Cacioppo & L.G. Tassinary (Eds.), *Principles of Psychophysiology*:

- Physical, Social, and Inferential Elements*. Cambridge: Cambridge University Press.
- Costen, N.P., Parker, D.M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics*, **58**, 602-612.
- Descharnes, R. (1972). *The World of Salvador Dali*. London: MacMillan London Limited.
- Deutsch, J.A., & Deutsch, A. (1963). Attention: Some theoretical considerations. *Psychological Review*, **70**, 80-90.
- DeValois, R.L., & DeValois, K.K. (1990). *Spatial Vision*. New York: Oxford University Press.
- Di Russo, F., & Spinelli, D. (2002). Effects of sustained, voluntary attention on amplitude and latency of steady-state visual evoked potential: a costs and benefits analysis. *Clinical Neurophysiology*, **113**, 1771-1777.
- Doallo, S., Lorenzo-López, L., Vizoso, C., Rodríguez Holguín, S., Amenedo, E., Bará, S., & Cadaveira, F. (2004). The time course of the effects of central and peripheral cues on visual processing: an event-related potentials study. *Clinical Neurophysiology*, **115**, 199-210.
- Dosher, B.A., & Lu, Z.L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 13988-13993.
- Dosher, B.A., & Lu, Z.L. (1999). Mechanisms of perceptual learning. *Vision Research*, **39**, 3197-3221.

- Doshier, B.A., & Lu, Z.-L. (2000a). Noise exclusion in spatial attention. *Psychological Science*, **11**, 139-146.
- Doshier, B.A., & Lu, Z.-L. (2000b). Mechanisms of perceptual attention in precuing of location. *Vision Research*, **40**, 1269-1292.
- Downing, C.J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 188-202.
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, **87**, 272-300.
- Dutta, A. (1995). Experimental run time system: software for developing and running experiments on IBM-compatible PCs. *Behavioural Research Methods, Instruments, & Computers*, **27**, 16-51.
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and clinical Neurophysiology*, **99**, 225-234.
- Eimer, M. (1998). Mechanisms of visuospatial attention: Evidence from event-related brain potentials. *Visual Cognition*, **5**, 257-286.
- Eimer, M. (2000). The time course of spatial orienting elicited by central and peripheral cues: Evidence from event-related brain potentials. *Biological Psychology*, **53**, 253-258.
- Eriksen, C.W., & St. James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom-lens model. *Perception & Psychophysics*, **40**, 225-240.
- Fahle, M. (2004). Perceptual learning: A case for early selection. *Journal of Vision*, **4**, 879-890, <http://journalofvision.org/4/10/4/>, DOI10.1167/4.10.4.

- Fahle, M., & Edelman, S. (1993). Long term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research*, **33**, 397-412.
- Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research*, **35**, 3003-3013.
- Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, **6**, 292-297.
- Fine, I., & Jacobs, R.A. (2000). Perceptual learning for a pattern discrimination task. *Vision Research*, **40**, 3209-3230.
- Fine, I., & Jacobs, R.A. (2002). Comparing perceptual learning across tasks: A review. *Journal of Vision*, **2**, 190-203, <http://journalofvision.org/2/2/5>, DOI10.1167/2.2.5.
- Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature*, **287**, 43-44.
- Fiorentini, A., Maffei, L., & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception*, **12**, 195-201.
- Fu, S., Fan, S., Chen, L., & Zhuo, Y. (2001). The attentional effects of peripheral cueing as revealed by two event-related potential studies. *Clinical Neurophysiology*, **112**, 172-185.
- Gauthier, I., Williams, P., Tarr, M.J., & Tanaka, J. (1998). Training "greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, **38**, 2401-2428.
- Goffaux, V., Jemel, B., Jacques, C., Rossion, B., & Schyns, P.G. (2003). ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cognitive Science*, **27**, 313-325.

- Gold, J., Bennett, P.J., & Sekuler, A.B. (1999a). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, **39**, 3537-3560.
- Gold, J., Bennett, P.J., & Sekuler, A.B. (1999b). Signal but not noise changes with perceptual learning. *Nature*, **402**, 176-178.
- Gold, J., Murray, R.F., Bennett, P.J., & Sekuler, A.B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology*, **10**, 663-666.
- Gold, J.M., Sekuler, A.B., & Bennett, P.J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, **28**, 167-207.
- Goldstone, R.L. (1998). Perceptual learning. *Annual Review of Psychology*, **49**, 585-612.
- Gosselin, F., Bacon, B., & Mamassian, P. (2004). Internal surface representations approximated by reverse correlation. *Vision Research*, **44**, 2515-2520.
- Gosselin, F., & Schyns, P.G. (2001). Bubbles: A new technique to reveal the use of visual information in recognition tasks. *Vision Research*, **41**, 2261-2271.
- Gosselin, F., & Schyns, P.G. (2003). Superstitious perceptions reveal properties of internal representations. *Psychological Science*, **14**, 505-509.
- Gosselin, F., & Schyns, P.G. (2004). No trouble with bubbles: a reply to Murray and Gold. *Vision Research*, **44**, 471-477.
- Hayes, T., Morrone, M.C., & Burr, D.C. (1986). Recognition of positive and negative bandpass-filtered images. *Perception*, **15**, 595-602.
- Hayes, A., & Ross, J. (1995). Lines of Sight. In R. Gregory, J. Harris, P. Heard & D. Rose, (Eds.), *The Artful Eye*. Oxford: Oxford University Press.

- Henderson, J.M. (1996). Spatial precues affect target discrimination in the absence of visual noise. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 780-787.
- Henning, G.B., Hertz, B.G., & Broadbent, D.E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research*, **15**, 887-897.
- Hillyard, S.A., Luck, S.J., & Mangun, G.R. (1994). The cuing of attention to visual field locations: Analysis with ERP recordings. In H. J. Heinze, T. F. Munte, & G. R. Mangun (Eds.), *Cognitive electrophysiology: Event-related brain potentials in basic and clinical research* (pp. 1-25). Boston: Birkhausen.
- Hillyard, S.A., Vogel, E.K., & Luck, S.J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London B*, **353**, 1257-1270.
- Hopfinger, J.B., & Mangun, G.R. (1998). Reflexive attention modulates processing of visual stimuli in human extrastriate cortex. *Psychological Science*, **9**, 441-447.
- Huyhn, H. (1978). Some approximate tests for repeated-measures designs. *Psychometrika*, **43**, 161-175.
- James, W. (1890). *The principles of psychology: Volume one*. New York: Dover Publications.
- Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 861-868.

- Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, **350**, 250-252.
- Legge, G., Kersten, D., & Burgess, A.E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America: A*, **4**, 391-406.
- Leopold, D.A., & Logothetis, N.K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, **3**, 254-264.
- Leuthold, H., Sommer, W., & Ulrich, R. (1996). Partial advance information and response preparation: Inferences from the lateralised readiness potential. *Journal of Experimental Psychology: General*, **125**, 307-323.
- Li, R.W., Levi, D.M., & Klein, S.A. (2004). Perceptual learning improves efficiency by re-tuning the decision 'template' for position discrimination. *Nature Neuroscience*, **7**, 178-183.
- Liebe, S., Gold, J.M., Busey, T.A., & O'Donnell, B. (2004). Electrophysiological correlates of the effects of perceptual learning on signal and noise in the human visual system. [Abstract]. *Journal of Vision*, **4**, 297a, <http://journalofvision.org/4/8/297/>, DOI:10.1167/4.8.297.
- Liu, Z. (1996). Viewpoint-dependency in object representation and recognition. *Spatial Vision*, **9**, 491-521.
- Liu, Z., Knill, D.C., & Kersten, D. (1995). Object classification for human and ideal observers. *Vision Research*, **35**, 549-568.
- Liu, Z., Knill, D.C., & Kersten, D. (1999). Dissociating stimulus information from internal representation – a case study on object recognition. *Vision Research*, **39**, 603-612.
- Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research*, **40**, 97-109.

- Livingstone, M. (2000). Is it warm? Is it real? Or just low spatial frequency? *Science*, **290**, 1299.
- Lu, Z.L., & Doshier, B.A. (1998). External noise distinguishes attention mechanisms. *Vision Research*, **38**, 1183-1198.
- Lu, Z.L., & Doshier, B.A. (2000). Spatial attention: Different mechanisms for central and peripheral temporal pre-cues? *Journal of Experimental Psychology : Human Perception and Performance*, **26**, 1534-1548.
- Lu, Z.L., & Doshier, B.A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision*, **4**, 44-56, <http://journofvision.org/4/1/5/>, DOI:10.1167/4.1.5.
- Lu, Z.L., Lesmes, L.A., & Doshier, B.A. (2002). Spatial attention excludes external noise at the target location. *Journal of Vision*, **2**, 312-323.
- Lubbe, R.H.J., & Woestenburger, J.C. (1997). Modulation of early ERP components with peripheral precues: A trend analysis. *Biological Psychology*, **45**, 143-158.
- Luck, S.J., & Hillyard, S.A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, **31**, 291-308.
- Luck, S.J., Woodman, G.F., & Vogel, E.K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, **4**, 432-440.
- Majaj, N.J., Pelli, D.G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, **42**, 1165-1184.
- Mangini, M.C., & Biederman, I. (2004). Making the ineffable explicit: estimating the information employed for face classifications. *Cognitive Science*, **28**, 209-226.

- Mangun, G.R.R., & Hillyard, S.A. (1987). The spatial allocation of visual attention as indexed by event-related brain potentials. *Human Factors*, **29**, 195-211.
- Maurer, D., Le Grand, R., & Mondloch, C.J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, **6**, 255-260.
- McCarthy, G., & Donchin, E. (1981). A metric for thought: A comparison of P300 latency and reaction time. *Science*, **211**, 77-80.
- McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction Searches. *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 1517-1539.
- Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, **4**, 539-551, <http://journofvision.org/4/7/2/>, DOI:10.1167/4.7.2.
- Miller, J., Patterson, T., & Ulrich, R. (1998). Jack-knife-based method for measuring LRP onset latency differences. *Psychophysiology*, **35**, 99-115.
- Morrisson, D., & Schyns, P.G. (2001). Usage of spatial scales for the categorization of faces, objects and scenes. *Psychological Bulletin and Review*, **8**, 454-469.
- Müller, M.M., & Hübner, R. (2002). Can the spotlight of attention be shaped like a donut? *Psychological Science*, **13**, 119-124.
- Müller, H.J., & Rabbitt, P.M.A. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 315-330.

- Murray, R.F., & Gold, J.M. (2004a). Troubles with bubbles. *Vision Research*, **44**, 461-470.
- Murray, R.F., & Gold, J.M. (2004b). Reply to Gosselin and Schyns. *Vision Research*, **44**, 479-482.
- Oliva, A. & Schyns, P.G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, **34**, 72-107.
- Osman, A., & Moore, C.M. (1993). The locus of dual-task interference: Psychological refractory effects on movement-related brain potentials. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 1-21.
- Pachella, R.G. (1974). The interpretation of reaction time in information-processing research. In B. Kantowitz (Ed.), *Human Information Processing: Tutorials in Performance and Cognition*. Hillsdale, NJ: Erlbaum.
- Pantle, A., & Sekuler, R. (1968). Size-detecting mechanisms in human vision. *Science*, **162**, 1146-1148.
- Parish, D.H., & Sperling, G. (1991). Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Research*, **31**, 1399-1415.
- Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- Pelli, D.G. (1981). *Effects of visual noise*. PhD Thesis, University of Cambridge, Cambridge.
- Pelli, D.G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 3-24). Cambridge: Cambridge University Press.
- Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.

- Pelli, P.G., & Farrell, B. (1999). Why use noise? *Journal of the Optical Society of America: A*, **16**, 647-653.
- Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science*, **256**, 1018-1021.
- Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **2**, 3-25.
- Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention and Performance X*. (pp.531-556). Hillsdale, N.J.:Erlbaum.
- Posner, M.I., Snyder, C.R.R., & Davidson, B.J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, **109**, 160-174.
- Pylyshyn, Z.W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioural and Brain Sciences*, **22**, 341-423.
- Rafal, R.D., Calabresi, P.A., Brennan, C.W., & Sciolto, T.K. (1989). Saccade preparation inhibits reorienting to recently attended locations. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 673-685.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59-108.
- Reed, A.V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, **181**, 574-576.
- Rensink, R.A., O'Regan, J.K., & Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, **8**, 368-373.

- Rensink, R.A. (2002). Change detection. *Annual Review of Psychology*, **53**, 245-277.
- Rinkenauer, G., Osman, A., Ulrich, R., & Mattes, S. (2004). On the locus of speed-accuracy tradeoff in reaction time: Inferences from the lateralised readiness potential. *Journal of Experimental Psychology: General*, **133**, 261-282.
- Rubin, G.S., & Siegel, K. (1984). Recognition of low-pass filtered faces and letters. *Investigative Ophthalmology and Visual Science*, **25**, 96.
- Rugg, M.D., & Coles, M.G.H. (1995). The ERP and cognitive psychology: Conceptual issues. In M.D. Rugg & M.G.H. Coles (Eds.), *Electrophysiology of Mind. Event-related Brain Potentials and Cognition* (pp. 27-39). Oxford: Oxford University Press.
- Schneider, K.A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, **47**, 333-366.
- Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information and their interactions. *Cognition*, **67**, 147-179.
- Schyns, P.G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, **13**, 402-409.
- Schyns, P.G., Jentzsch, I., Johnson, M., Schweinberger, S.R., & Gosselin, F. (2003). A principled method for determining the functionality of brain responses. *NeuroReport*, **14**, 1665-1669.
- Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, **69**, 243-265.

- Schyns, P.G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **23**, 681-696.
- Sekuler, R., & Blake, R. (1994). *Perception*. New York: McGraw-Hill.
- Shiu, L., & Pashler, H. (1995). Spatial attention and vernier acuity. *Vision Research*, **35**, 337-343.
- Simoncelli, E.P. (1997). Image and multi-scale pyramid tools [Computer software]. NewYork: Author.
- Simons, D.J., & Rensink, R.A. (2005). Change blindness: Past, present and future. *Trends in Cognitive Sciences*, **9**, 16-20.
- Smith, M.L., Cottrell, G.W., Gosselin, F., & Schyns, P.G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, **16**, 184-189.
- Smith, P.L. (2000). Attention and luminance detection: Effects of cues, masks, and pedestals. *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 1401-1420.
- Solomon, J.A., & Pelli, D.G. (1994). The visual filter mediating letter identification. *Nature*, **369**, 395-397.
- Sowden, P.T., Davies, I.R.L., & Roling, P. (2000). Perceptual learning of the detection of features in x-ray images: A functional role for improvements in adults' visual sensitivity? *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 379-390.
- Sowden, P.T., Davies, I., Rose, D., & Kaye, M. (1996). Perceptual learning of stereoacuity. *Perception*, **25**, 1043-1052.
- Sowden, P.T., Özgen, E., Schyns, P.G., & Daoutis, C. (2003). Expectancy effects on spatial frequency processing. *Vision Research*, **43**, 2759-2772.

- Stromeyer, C.F., & Klein, S. (1975). Evidence against narrow-band spatial frequency channels in human vision: The detectability of frequency modulated gratings. *Vision Research*, **15**, 899-910.
- Talgar, C.P., & Carrasco, M. (2002). Vertical meridian asymmetry in spatial resolution: Visual and attentional factors. *Psychonomic Bulletin & Review*, **9**, 714-722.
- Tarr, M.J., & Bülthoff, H.H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 1494-1505.
- Tieger, T., & Ganz, L. (1979). Recognition of faces in the presence of two-dimensional sinusoidal masks. *Perception and Psychophysics*, **26**, 163-167.
- Tjan, B.S., Braje, W.L., Legge, G.E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, **35**, 3053-3069.
- Tjan, B.S., Chung, T.S.L., & Levi, D.M. (2002a). Limitation of ideal-observer analysis in understanding perceptual learning. Annual Meeting Abstract and Program Planner accessed at www.arvo.org. Association for Research in Vision and Ophthalmology. Abstract 2916.
- Tjan, B.S., Chung, T.S.L., & Levi, D.M. (2002b). How many functional factors does it take to explain perceptual learning? *Perception (supp)*, **31**, 109a.
- Tjan, B.S., & Legge, G.E. (1998). The viewpoint complexity of an object-recognition task. *Vision Research*, **38**, 2335-2350.
- Tjan, B.S., Lestou, V., Bulthoff, H.H., & Kourtzi, Z. (2003). An fMRI method for identifying the sequential stages of processing in the ventral visual pathway. *Journal of Vision*, **3**, 109a, <http://journalofvision.org/3/9/10> doi:10.1167/3.9.109.

- Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, **34**, 131-156.
- Vinette, C., Gosselin, F., & Schyns, P.G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science*, **28**, 289-301.
- Wallace, J.M., & Mamassian, P. (2003). The efficiency of speed discrimination for coherent and transparent motion. *Vision Research*, **43**, 2795-2810.
- Watanabe, T., Náñez, J.E., Koyama, S., Mukai, I., Liederman, J., & Sasaki, Y. (2002). Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nature Neuroscience*, **5**, 1003-1009.
- Watanabe, T., Náñez, J.E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, **413**, 844-848.
- Webster, M.A., & Miyahara, E. (1997). Contrast adaptation and the spatial structure of natural images. *Journal of the Optical Society of America A*, **14**, 2355-2366.
- Wickelgren, W.A. (1977). Speed-accuracy trade-off and information processing dynamics. *Acta Psychologica*, **41**, 67-85.
- Wilson, H.R., & Bergen, J.R. (1979). A four mechanism model for threshold spatial vision. *Vision Research*, **19**, 19-32.
- Wilson, H.R., McFarlane, D.K., & Phillips, G.C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, **23**, 873-882.
- Wilson, H.R., & Wilkinson, F. (1997). Evolving concepts of spatial channels in vision: From independence to nonlinear interactions. *Perception*, **26**, 939-960.
- Yarbus, A.L. (1965). *Eye movements and vision*. New York: Plenum.

- Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, **396**, 72-75.
- Yeshurun, Y., & Carrasco, M. (2000). The locus of attentional effects in texture segmentation. *Nature Neuroscience*, **3**, 622-627.